ELECTRON-PROTON NONADIABATICITY: CHARACTERIZATION AND DEVELOPMENT OF NON-BORN-OPPENHEIMER ELECTRONIC STRUCTURE METHODS

BY

ANDREW RAJENDRA SIRJOOSINGH

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Chemistry
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Doctoral Committee:

Professor Sharon Hammes-Schiffer, Chair
Professor So Hirata
Professor Nancy Makri
Professor David Ceperley

ABSTRACT

Nuclear quantum effects such as zero-point energy and hydrogen tunneling play an important role in a wide variety of chemical reactions. Moreover, non-Born-Oppenheimer effects are important in reactions such as proton-coupled electron transfer (PCET), which are integral to various electrocatalytic applications and bioenzymatic processes. The breakdown of the Born-Oppenheimer approximation between electronic and nuclear motions engenders the need for accurate characterization of the degree of nonadiabaticity. Furthermore, in regimes where the inclusion of these effects is vital, as it is for PCET systems, the development of non-Born-Oppenheimer quantum chemical methods is increasingly important. In this dissertation, we present diagnostics of electron-proton nonadiabaticity that can be obtained from standard electronic structure calculations and describe their application to representative systems, highlighting the mechanistic differences between two subclasses of PCET. In addition, we describe the development of new electronic structure methods within the nuclear-electronic orbital (NEO) framework, which is an orbital-based approach that inherently includes electron-proton nonadiabaticity by treating electrons and select protons quantum mechanically on equal footing. Previous studies using NEO involved applying mean-field-based approaches, which lacked sufficient electron-proton dynamical correlation, leading to overlocalized nuclear densities. Subsequent efforts focused on the development of explicitly correlated NEO approaches which, although accurate, were too computationally intractable to be practical for the study of PCET systems. In this dissertation, we describe two approaches to develop tractable NEO methods. Firstly, we describe the formulation of a multi-component density functional theory approach within the NEO framework, which involves the derivation of several electron-

proton correlation functionals to accurately account for electron-proton correlation. Secondly, we describe in detail a novel NEO method: the reduced explicitly correlated Hartree-Fock (RXCHF) approach, which is a wavefunction-based approach that accurately accounts for electron-proton dynamical correlation between a subset of the electronic orbitals and the quantum nuclear orbitals. Systematic approximations for the RXCHF methods that afford substantial gains in computational tractability will be described, and model calculations demonstrating the applicability of this method to systems with select protons treated quantum mechanically will be presented. This method will enable accurate calculations on PCET systems as electron-proton nonadiabatic effects are inherently included, and it can provide fundamental insight into these types of mechanisms. In addition, the NEO-RXCHF method readily affords a framework within which experimentally accessible quantities such as rate constants and isotope effects can be determined for PCET reactions.

ACKNOWLEDGEMENTS

The path toward the completion of my doctoral dissertation has been very rewarding. I have benefited from superb intellectual guidance from members of our research group, and have been extremely well supported throughout my tenure by friends and family.

I first must thank my advisor, Professor Sharon Hammes-Schiffer, for her outstanding guidance during my doctoral studies. I have learned a great deal about science thanks to her expertise and will always be grateful for the time and patience shown to me, especially when I was a younger graduate student just starting in the group. I thank her for allowing me the opportunity to work on numerous interesting projects over the years, and for continually encouraging me to explore new research directions inspired by my own work. I have benefited significantly from being able to work closely with such an accomplished scientist and will always be thankful for having such an opportunity for the duration of my doctoral work.

I also thank all the members of the Hammes-Schiffer research group with whom I have crossed paths during my doctoral studies. I have benefited from interactions with many members over the years, and I must directly thank the collaborators on the NEO project. I have greatly enjoyed working with Drs. Chaehyuk Ko, Chet Swalina and Mike Pak, whose patience and guidance have all been invaluable. In particular, I thank Dr. Mike Pak for his outstanding scientific guidance throughout the entirety of my graduate studies. His scientific prowess has been a great example and I am grateful to have had the opportunity to benefit from working closely with him. I also thank Dr. Alexander Soudackov who, although I have not worked with directly, has offered a great deal of help to me and, as most group members would agree, represents a paragon of scientific excellence. His support and knowledge have also been

invaluable to me over the duration of my studies. Finally, I am grateful for having two great colleagues who began graduate studies in the Hammes-Schiffer group at the same time as me: Abir Ganguly and Brian Solis, who have shared in classes, exams and dissertation writing throughout the duration of our graduate studies.

Any success I can derive from my doctoral studies could not have been possible without the great deal of support I have received from friends and family. I first must thank my dear fiancée, Pallavi Thaplyal, who I was fortunate enough to meet in graduate school. I have never encountered someone as supportive and as caring as her, and I will be forever grateful for her patience, kindness and, not in the least, her scientific aptitude. I am extremely thankful to have shared this journey together with her, and without her support, I would not nearly have been as successful as I might call myself today.

In addition, I thank my parents, Cecil and Chan Sirjoosingh, who have both provided a huge amount of support to me over the years, and have always encouraged me to strive for excellence. Their support throughout my life has enabled this achievement, and for that I extend to them my sincerest gratitude. I also thank my sister, Candace Sheppard, and my brother-in-law, Dr. Kenneth Sheppard, who have provided excellent examples for me in my pursuit of an advanced academic degree. Their dedication to their own studies has been a significant inspiration to me during the course of my graduate tenure, and I will always be grateful for their support over the years. I always look forward to visiting with them and the four of us, with Pallavi, will hopefully celebrate our academic successes in Scotland in the near future!

TABLE OF CONTENTS

Chapter 1

Introduction

1.1. Proton-coupled electron transfer

Processes involving the coupled transfer of electrons and protons occur readily in nature and are important to the study of a wide variety of chemical and biological systems.[1-5] These processes, often termed proton-coupled electron transfer (PCET) reactions, have been widely studied over the past few decades using both experimental and theoretical approaches. A schematic of a PCET process is given in Figure 1.1, where a single electron and proton transfer from their respective donors to their respective acceptors. In general, the donors may be the same or distinct for the electron and proton, and likewise for the acceptors.

Some important applications of PCET reactions include vital energy conversion and storage processes. The most familiar of these is likely photosynthesis, which involves the transfer of several protons coupled to several electron transfer steps resulting in the conversion of carbon dioxide and water to carbohydrates and oxygen. PCET also plays an integral role in various electrochemical processes including the catalytic oxidation or reduction of hydrogen via molecular electrocatalysts. Synopses of various electrochemical methods that exploit PCET for some of these types of processes are given in some comprehensive reviews and the references therein.[5-10]

Additional applications of PCET processes can be found in biological systems, including enzymatic processes involving proteins and even in nucleic acid chemistry. Several recent reviews have highlighted progress in understanding these types of processes in various enzymes, such as oxidases and hemes, as well as various biomimetic models.[11-17] With the development of sophisticated theoretical and experimental techniques, insight into these types of systems is becoming increasingly available.

Another class of applications of PCET can be found in various photoinduced processes. The initiation of PCET reactions by photoexcitation affords interesting experimental and theoretical challenges, as the involvement of many excited electronic and vibronic states can impact the dynamics of the resulting reactions, and many relaxation pathways could potentially be involved. Interesting photoinduced PCET processes are currently being studied in DNA as well as in other applications.[5,18-22]

Theoretical models developed for PCET mechanisms have been prevalent since the early experimental work on these types of systems.[1] Since their inception, a main consideration has been the nature of the coupling between the electrons and protons in a PCET process. Established rate theories for PCET processes have been developed over the past two decades in the Hammes-Schiffer group and applied to a variety of different PCET applications, including condensed phase processes, electrochemical applications, biological systems and most recently, photoinduced processes. A detailed analysis of these developments can be found in recent reviews on this and other theoretical approaches.[2,14,23-27] In what follows, we present an overview of these theories, highlighting some of the important underlying considerations concerning the nature of the coupling between electron and proton motion.

The rate theories developed for PCET in the Hammes-Schiffer group apply an extension of the Marcus theory[28] approach for electron transfer.[26,29-31] In particular, rate constant expressions were derived in the vibronically nonadiabatic limit, where it is assumed that the electrons and transferring proton do not respond instantaneously to the solvent motion. In this regime, Fermi's Golden rule formalism may be applied to obtain the following rate constant expression,[26]

$$k = \sum_{\mu} P_{\mu} \sum_{\nu} \frac{\left|V_{\mu\nu}\right|^2}{\hbar} \sqrt{\frac{\pi}{\lambda_{\mu\nu} k_B T}} \exp\left(-\frac{\left(\Delta G^{\circ}_{\mu\nu} + \lambda_{\mu\nu}\right)^2}{4\lambda_{\mu\nu} k_B T} \right), \tag{1.1}$$

where the summation indices μ and ν are over reactant and product vibronic states, respectively, P_{μ} is the Boltzmann population of reactant vibronic state μ, $V_{\mu\nu}$ is the vibronic coupling between μ and ν, $\lambda_{\mu\nu}$ is the reorganization energy associated with μ and ν, $\Delta G^{\circ}_{\mu\nu}$ is the free energy of reaction for states μ and ν, k_B is the Boltzmann constant and T is the temperature. This expression is analogous to the standard Marcus theory expression, where the rate constant of electron transfer (ET) is proportional to the square of the electronic coupling, although here, the rate constant of PCET is proportional to the square of the vibronic coupling, $V_{\mu\nu}$. Some of these quantities are graphically represented in Figure 1.2.

Within the vibronically nonadiabatic limit where Eq. (1.1) is valid, we may still consider two limiting regimes concerning the nature of the coupling between electron and proton motion. In the electronically adiabatic limit, the electrons respond instantaneously to the motion of the transferring proton, whereas in the electronically nonadiabatic limit, they do not. The form of the vibronic coupling is different in both of these two regimes, and thus the relative degree of electron-proton nonadiabaticity ultimately impacts the form of the rate constant expression in Eq.

(1.1).[25,30,32,33] This dependence will, in turn, affect the form of the calculated kinetic isotope effects, which are ratios of rate constants of the PCET processes using protium versus deuterium. General expressions for the vibronic coupling applicable in both regimes have recently been derived using a semiclassical approach,[32,33] but knowledge of the relative amount of electron-proton nonadiabaticity can provide additional insight into the PCET mechanism, as will be discussed in this dissertation.

Another consideration of applying the rate theories previously described is the definition of reactant and product vibronic states. Analogous to Marcus theory for electron transfer, these reactant and product vibronic states correspond to the transferring electron and proton localized on their donors or their acceptors, respectively. Previously, these states were fit to empirical models built for specific applications,[26,33] but since most quantities appearing in Eq. (1.1) are defined in terms of these states, it would be beneficial to devise a general procedure to determine these localized states. Such approaches will be addressed in this dissertation.

Further understanding of electron-proton nonadiabaticity, which is a key consideration in the application of PCET rate theories, can serve to improve the understanding of PCET processes in order to provide insight into the mechanism. In addition, general methods to define the input vibronic state quantities will enable accurate and streamlined calculations on a wide range of PCET applications. These two points will be addressed in the first chapters of this dissertation, and the work presented herein can be used in the application of PCET rate theories to enable accurate prediction of experimentally relevant quantities such as rate constants and kinetic isotope effects as well as to gain further insight into PCET mechanisms.

1.2. Non-Born-Oppenheimer Methods

The notion of electron-proton nonadiabatic effects alluded to in the previous subsection can be more formally defined with relation to the Born-Oppenheimer (B-O) approximation.[34] When applying the B-O approximation to standard problems in quantum chemistry, one is inherently assuming that the electrons move much faster than the nuclei, and thus the former are treated using quantum mechanics and the latter using classical mechanics. This approximation is what allows the identification and manipulation of chemical structures, where it is implicitly assumed that the electrons are delocalized around a fixed configuration of classical nuclei. For dynamical treatments using the B-O approximation, one usually invokes the analogous assumption that the electrons respond instantaneously to the motion of the nuclei, and thus the motion of the nuclei can be considered as evolving on the ground electronic adiabatic state (i.e., the potential energy surface computed by calculating electronic energies as a function of the various nuclear positions). Specifically, at each configuration of the nuclei, the time-independent Schrödinger equation is solved for the electronic Hamiltonian given by

$$H = T_e + V_{ec} + V_{ee}, \tag{1.2}$$

where T_e is the kinetic energy operator for the electrons, V_{ec} is the Coulomb attraction between the electrons and the fixed configuration of classical nuclei and V_{ee} is the Coulomb repulsion between electrons.

When the B-O approximation breaks down, so-called non-Born-Oppenheimer (non B-O), or nonadiabatic effects are significant. In these cases, the motion of the electrons cannot be interpreted as much faster than the motion of all nuclei, and in order to effectively describe nuclear motion, excited electronic states must be considered. One particular class of systems where the B-O approximation is often inadequate is PCET, where in particular, the B-O

separation between the electronic motion and the motion of the transferring proton is invalid. Thus, non B-O effects between the electrons and the transferring proton, or electron-proton nonadiabatic effects, are important to consider when applying theoretical techniques to the study of PCET systems.

Standard B-O electronic structure methods have been developed, optimized and applied over numerous decades.[35-37] The ever-increasing computational capabilities have facilitated the regular usage of many of these approaches, and their increased prevalence has now made it commonplace to perform very accurate quantum chemical calculations on smaller molecular systems, and reasonably accurate calculations (enough to draw qualitative insights commensurate with experimental hypotheses) on larger chemical systems. Applying these methods to PCET systems, however, is not always appropriate due to the importance of the electron-proton nonadiabatic effects previously discussed.

The development of non B-O electronic structure approaches has been significantly less prevalent than that of the aforementioned B-O methods due mainly to the long-lasting success of B-O methods to straightforward problems in quantum chemistry and the additional computational expense that is incurred by the inclusion of nonadiabatic effects. Dynamical treatments that go beyond the B-O approximation include semiclassical treatments as well as path integral methods. A few important approaches as well as recent review articles can be found in Refs. 38-46. These methods aim in part to solve the time-dependent Schrödinger equation for the nuclei moving on electronic potential energy surfaces often calculated on the fly. Within the context of PCET mechanisms, however, it is desirable to study the stationary states of a mixed nuclear-electronic subsystem in the presence of an external potential by solving the

time-independent Schrödinger equation for the particles comprising this subsystem (usually the electrons and select nuclei treated quantum mechanically).

For PCET processes, typically the electrons and the transferring proton comprise the subsystem to be treated quantum mechanically, while all other nuclei in the system are treated classically. A given configuration of the classical nuclei then forms a fixed external potential for which the corresponding time-independent Schrödinger equation can be solved for the quantum subsystem. The associated mixed nuclear-electronic Hamiltonian is given by

$$H = T_e + T_p + V_{ec} + V_{pc} + V_{ee} + \mathrm{V}_{pp} + \mathrm{V}_{ep}, \tag{1.3}$$

where augmenting the terms present in the electronic Hamiltonian of Eq. (1.2) are the kinetic energy of the quantum nuclei, T_p, the Coulomb repulsion of the quantum nuclei with the classical nuclei, V_{pc}, the Coulomb repulsion of the quantum nuclei between themselves, V_{pp}, and the Coulomb attraction between electrons and the quantum nuclei, V_{ep}. This type of approach inherently includes nonadiabatic effects between the quantities comprising the quantum subsystem, and thus, when applied to PCET, includes the electron-proton nonadiabatic effects which are often important. As a result, quantities routinely calculated using B-O approaches for reactions including minimum energy paths, geometry optimized structures and vibrational frequencies can be calculated with the relevant nonadiabatic effects implicitly included.

One such non B-O approach that aims to solve the time-independent Schrödinger equation corresponding to the Hamiltonian in Eq. (1.3) is the nuclear-electronic orbital (NEO) approach, where electrons and select nuclei are treated quantum mechanically using orbital-based techniques analogous to standard electronic structure methods.[47-53] Similar methods are also being developed by other groups.[54-67] In the NEO approach, the select quantum nuclei (e.g., the transferring proton(s) in a PCET reaction) are provided with an atomic basis set analogous to

that used for the electrons in conventional B-O electronic structure theory, and after an ansatz for the mixed nuclear-electronic wavefunction is proposed, the energy is computed by evaluating matrix elements over the Hamiltonian corresponding to the quantum subsystem. Variations of the energy with respect to molecular orbitals for the electrons as well as the nuclei built up from their respective atomic orbital basis sets results in a variational self-consistent-field procedure that can be used to minimize the energy, providing an optimal mixed nuclear-electronic wavefunction corresponding to a stationary state of the system. A pictorial example of the quantization of a hydrogen nucleus can be found in Figure 1.3 for the hydrogen cyanide molecule.

A main conclusion of previous NEO work has been that an accurate description of electron-nuclear correlation is important to capture even a qualitatively accurate picture, mainly due to the attractive interaction between electrons and the quantum nuclei.[51,68-70] To this end, explicitly correlated NEO methods were developed,[48,49,51] and although significantly more accurate than their mean-field counterparts, they proved to be too costly computationally. In this dissertation, we discuss progress toward developing novel explicitly correlated density functional theory (DFT) and wavefunction-based approaches within the NEO framework that, in addition to providing a sufficient level of accuracy, still maintain a reasonable level of tractability. The tradeoff between accuracy and tractability had previously plagued explicitly correlated NEO approaches, but in what follows, we present some new approaches where, through systematic approximations, calculations can be made significantly more efficient while preserving a suitable level of accuracy. The advent of tractable, yet accurate non B-O methods within the NEO framework offers exciting possibilities for the study of PCET systems. In particular, these methods will provide increased insight into the mechanisms of these processes in addition to

enabling the accurate computation of key quantities required for calculating experimentally relevant quantities such as rate constants and kinetic isotope effects.

1.3. Outline

In the remainder of this dissertation, we will elaborate on the notion of electron-proton nonadiabaticity and its relevance to PCET reactions, and describe the development of new non-Born-Oppenheimer electronic structure methods that are applicable to situations where nonadiabatic effects are important.

In Chapter 2, we discuss our work on characterizing electron-proton nonadiabaticity and generating charge-localized diabatic electronic states for PCET reactions. In particular, we examine two prototypical PCET systems, the phenoxyl-phenol and benzyl-toluene radical couples, which represent two regimes of adiabaticity. We develop both qualitative and quantitative diagnostics of electron-proton nonadiabaticity and apply them to these two systems to demonstrate the applicability of this characterization scheme to PCET systems. As part of this study, we also develop a scheme to generate charge-localized diabatic electronic states for PCET reactions, which can serve as the input quantities for PCET rate theories.

Chapter 3 extends the diabatization work introduced in Chapter 2. In particular, we develop alternative formulations of the scheme used to generate charge-localized diabatic electronic states in order to perform this procedure along different types of reaction coordinates involving the proton motion. In addition, we discuss strategies to generate charge-localized diabatic electron-proton vibronic states corresponding to both electrons and the proton localized on donors or acceptors, which serve as the input quantities for existing PCET rate theories.

A common theme of the first two chapters is that electron-proton nonadiabatic effects are often prevalent in PCET reactions and are thus an important consideration in theoretical

approaches to studying these types of systems. Therefore, in subsequent chapters, we discuss the development of *ab initio* electronic structure methods within the nuclear-electronic orbital (NEO) framework that account for these types of effects. Chapters 4 and 5 discuss the implementation of multicomponent density functional theory-based methods, while Chapters 6, 7 and 8 describe the formulation of wavefunction-based approaches.

In Chapter 4, we derive electron-proton correlation functionals for use within the NEO-DFT method. We define these functionals in terms of key quantities arising from previously developed explicitly correlated NEO wavefunctions and demonstrate their applicability to several model systems.

Chapter 5 describes straightforward strategies to improve the previously developed electron-proton correlation functionals by the inclusion of kinetic energy terms involving both the electrons and the proton. In addition, we examine the interplay between electron-electron and electron-proton correlation and determine that these effects are predominantly additive, which supports the independent development of each type of correlation functional without the need for any reparameterization when both are used together.

We discuss the formulation of a new explicitly correlated wavefunction-based approach within the NEO framework in Chapter 6. This method, termed NEO-RXCHF (reduced explicitly correlated Hartree-Fock), aims to improve the tractability of previous explicitly correlated NEO methods by making different physical assumptions about the interactions between electrons and protons in PCET systems. In particular, this approach selects one electronic orbital to explicitly correlate to the quantum proton orbital. We develop this new method from first principles and discuss the key approximations that are necessary to enable practical application.

We apply the NEO-RXCHF method to several positron-containing systems in Chapter 7 in order to benchmark this approach as well as to demonstrate the validity of its inherent approximations. We also highlight its tractability advantages over previous explicitly correlated NEO methods.

Chapter 8 extends the NEO-RXCHF method to a situation where multiple electronic orbitals can be explicitly correlated to the quantum proton orbital. This extension was necessary in order to apply this method to the study of proton-containing systems. In this chapter, we present the application of NEO-RXCHF to molecular systems containing hydrogen; the first such calculations done on systems of this size within the NEO framework.

In Chapter 9, we present the conclusions of this dissertation and set the framework for current and future research in this area. In particular, the development of the non-Born-Oppenheimer electronic structure methods presented in this dissertation enables accurate calculations on systems where nonadiabatic effects are important, such as PCET systems. Continued development of these methods will ultimately lead to accurate and tractable methods which may be used to gain fundamental insight into the mechanisms of PCET reactions, as well enable the prediction of experimentally accessible quantities such as rate constants and kinetic isotope effects.

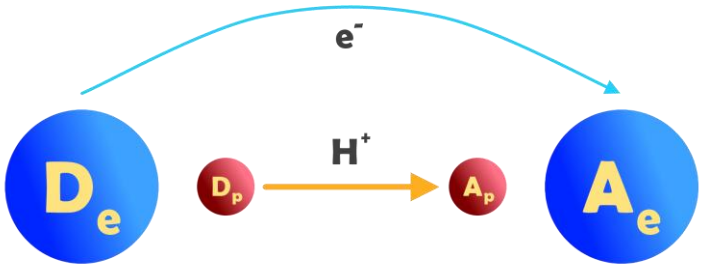

Figure 1.1: A schematic of a proton-coupled electron transfer reaction. The blue spheres, D_e and A_e, are the electron donor and acceptor, respectively, while the red spheres, D_p and A_p, are the proton donor and acceptor, respectively.

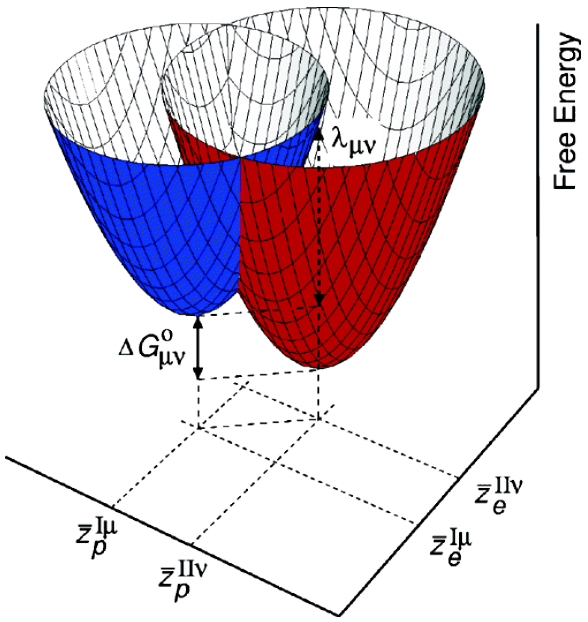

Figure 1.2: A schematic of free energy surfaces for reactant, μ (blue) and product, ν (red) vibronic states as a function of two collective solvent polarization coordinates highlighting the quantities discussed in the text. Reprinted with permission from Ref. 12. Copyright 2009 American Chemical Society.

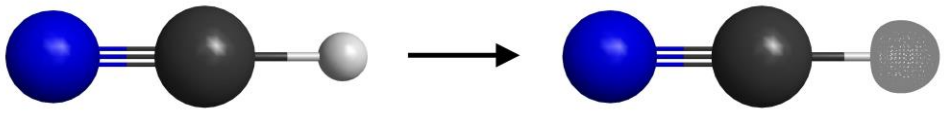

Figure 1.3: A schematic of the effect of quantization on the proton in the HCN molecule. On the left is a standard ball-and-stick model of HCN where the H is treated as a classical nucleus (hard sphere). On the right is the analogous pictorial representation of HCN when the H is quantized, where the ground state nuclear molecular orbital related to the probability distribution of the H position replaces the classical hard sphere.

References

(1) R. I. Cukier and D. G. Nocera, *Annual Review of Physical Chemistry* **49**, 337 (1998).
(2) S. Hammes-Schiffer, *Chemical Reviews* **110**, 6937 (2010).
(3) M. H. V. Huynh and T. J. Meyer, *Chemical Reviews* **107**, 5004 (2007).
(4) J. M. Mayer, *Annual Review of Physical Chemistry* **55**, 363 (2004).
(5) D. R. Weinberg, C. J. Gagliardi, J. F. Hull, C. F. Murphy, C. A. Kent, B. C. Westlake, A. Paul, D. H. Ess, D. G. McCafferty, and T. J. Meyer, *Chemical Reviews* **112**, 4016 (2012).
(6) C. Costentin, *Chemical Reviews* **108**, 2145 (2008).
(7) C. Costentin, C. Louault, M. Robert, and J. M. Saveant, *Proceedings of the National Academy of Sciences of the United States of America* **106**, 18143 (2009).
(8) C. Costentin, M. Robert, and J. M. Saveant, *Accounts of Chemical Research* **43**, 1019 (2010).
(9) J. L. Dempsey, B. S. Brunschwig, J. R. Winkler, and H. B. Gray, *Accounts of Chemical Research* **42**, 1995 (2009).
(10) M. R. DuBois and D. L. DuBois, *Chemical Society Reviews* **38**, 62 (2009).
(11) B. J. Bahnson, T. D. Colby, J. K. Chin, B. M. Goldstein, and J. P. Klinman, *Proceedings of the National Academy of Sciences of the United States of America* **94**, 12797 (1997).
(12) S. Hammes-Schiffer, *Accounts of Chemical Research* **39**, 93 (2006).
(13) S. Hammes-Schiffer and S. J. Benkovic, *Annual Review of Biochemistry* **75**, 519 (2006).
(14) J. P. Layfield and S. Hammes-Schiffer, *Chemical Reviews* **114**, 3466 (2014).
(15) Z. D. Nagel and J. P. Klinman, *Chemical Reviews* **106**, 3095 (2006).
(16) Z. D. Nagel and J. P. Klinman, *Nature Chemical Biology* **5**, 543 (2009).
(17) V. C. Nashine, S. Hammes-Schiffer, and S. J. Benkovic, *Current Opinion in Chemical Biology* **14**, 644 (2010).
(18) K. de La Harpe, C. E. Crespo-Hernandez, and B. Kohler, *Journal of the American Chemical Society* **131**, 17557 (2009).
(19) C. J. Gagliardi, B. C. Westlake, C. A. Kent, J. J. Paul, J. M. Papanikolas, and T. J. Meyer, *Coordination Chemistry Reviews* **254**, 2459 (2010).
(20) A. Kumar and M. D. Sevilla, *Chemical Reviews* **110**, 7002 (2010).
(21) B. Li, J. Zhao, K. Onda, K. D. Jordan, J. L. Yang, and H. Petek, *Science* **311**, 1436 (2006).
(22) O. V. Prezhdo, W. R. Duncant, and V. V. Prezhdo, *Accounts of Chemical Research* **41**, 339 (2008).
(23) S. Hammes-Schiffer, *Accounts of Chemical Research* **34**, 273 (2001).
(24) S. Hammes-Schiffer, *Accounts of Chemical Research* **42**, 1881 (2009).
(25) S. Hammes-Schiffer, *Energy & Environmental Science* **5**, 7696 (2012).
(26) S. Hammes-Schiffer and A. V. Soudackov, *Journal of Physical Chemistry B* **112**, 14108 (2008).
(27) S. Hammes-Schiffer and A. A. Stuchebrukhov, *Chemical Reviews* **110**, 6939 (2010).
(28) R. A. Marcus and N. Sutin, *Biochimica Et Biophysica Acta* **811**, 265 (1985).
(29) A. Soudackov and S. Hammes-Schiffer, *Journal of the American Chemical Society* **121**, 10598 (1999).
(30) A. Soudackov and S. Hammes-Schiffer, *Journal of Chemical Physics* **113**, 2385 (2000).
(31) A. Soudackov, E. Hatcher, and S. Hammes-Schiffer, *Journal of Chemical Physics* **122**, 014505 (2005).

(32) Y. Georgievskii and A. A. Stuchebrukhov, *Journal of Chemical Physics* **113**, 10438 (2000).
(33) J. H. Skone, A. V. Soudackov, and S. Hammes-Schiffer, *Journal of the American Chemical Society* **128**, 16655 (2006).
(34) M. Born and R. Oppenheimer, *Annalen Der Physik* **84**, 0457 (1927).
(35) A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. (Dover, New York, 1996).
(36) G. C. Schatz and M. A. Ratner, *Quantum Mechanics in Chemistry*. (Dover, Mineola, 2002).
(37) R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*. (Oxford University Press, New York, 1989).
(38) M. Ben-Nun, J. Quenneville, and T. J. Martinez, *Journal of Physical Chemistry A* **104**, 5161 (2000).
(39) D. F. Coker and L. Xiao, *Journal of Chemical Physics* **102**, 496 (1995).
(40) S. Habershon, D. E. Manolopoulos, T. E. Markland, and T. F. Miller, *Annual Review of Physical Chemistry* **64**, 387 (2013).
(41) S. Hammes-Schiffer and J. C. Tully, *Journal of Chemical Physics* **101**, 4657 (1994).
(42) R. Kapral, *Annual Review of Physical Chemistry* **57**, 129 (2006).
(43) N. Makri, *Annual Review of Physical Chemistry* **50**, 167 (1999).
(44) M. Topaler and N. Makri, *Journal of Chemical Physics* **101**, 7500 (1994).
(45) J. C. Tully, *Journal of Chemical Physics* **93**, 1061 (1990).
(46) G. A. Voth, *Advances in Chemical Physics* **93**, 135 (1996).
(47) A. Chakraborty, M. V. Pak, and S. Hammes-Schiffer, *Physical Review Letters* **101**, 153001 (2008).
(48) A. Chakraborty, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Physics* **129**, 014101 (2008).
(49) C. Ko, M. V. Pak, C. Swalina, and S. Hammes-Schiffer, *Journal of Chemical Physics* **135**, 054106 (2011).
(50) J. H. Skone, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Physics* **123**, 134108 (2005).
(51) C. Swalina, M. V. Pak, A. Chakraborty, and S. Hammes-Schiffer, *Journal of Physical Chemistry A* **110**, 9983 (2006).
(52) C. Swalina, M. V. Pak, and S. Hammes-Schiffer, *Chemical Physics Letters* **404**, 394 (2005).
(53) S. P. Webb, T. Iordanov, and S. Hammes-Schiffer, *Journal of Chemical Physics* **117**, 4106 (2002).
(54) A. D. Bochevarov, E. F. Valeev, and C. D. Sherrill, *Molecular Physics* **102**, 111 (2004).
(55) S. Bubin, M. Pavanelo, W. C. Tung, K. L. Sharkey, and L. Adamowicz, *Chemical Reviews* **113**, 36 (2013).
(56) M. Cafiero, S. Bubin, and L. Adamowicz, *Physical Chemistry Chemical Physics* **5**, 1491 (2003).
(57) M. Hoshino, H. Nishizawa, and H. Nakai, *Journal of Chemical Physics* **135**, 024111 (2011).
(58) E. Kamarchik and D. A. Mazziotti, *Physical Review A* **75**, 013203 (2007).
(59) E. Kamarchik and D. A. Mazziotti, *Physical Review A* **79**, 012502 (2009).
(60) T. Kreibich and E. K. U. Gross, *Physical Review Letters* **86**, 2984 (2001).

(61) T. Kreibich, R. van Leeuwen, and E. K. U. Gross, *Physical Review A* **78**, 022501 (2008).
(62) E. Matyus and M. Reiher, *Journal of Chemical Physics* **137**, 024104 (2012).
(63) H. Nakai, *International Journal of Quantum Chemistry* **86**, 511 (2002).
(64) H. Nakai, K. Sodeyama, and M. Hoshino, *Chemical Physics Letters* **345**, 118 (2001).
(65) Y. Shigeta, H. Nagao, K. Nishikawa, and K. Yamaguchi, *Journal of Chemical Physics* **111**, 6171 (1999).
(66) T. Shimazaki and M. Kubo, *Chemical Physics Letters* **525-26**, 134 (2012).
(67) M. Tachikawa, K. Mori, H. Nakai, and K. Iguchi, *Chemical Physics Letters* **290**, 437 (1998).
(68) M. V. Pak and S. Hammes-Schiffer, *Physical Review Letters* **92**, 103002 (2004).
(69) M. V. Pak, C. Swalina, S. P. Webb, and S. Hammes-Schiffer, *Chemical Physics* **304**, 227 (2004).
(70) C. Swalina and S. Hammes-Schiffer, *Journal of Physical Chemistry A* **109**, 10410 (2005).

Chapter 2

Proton-Coupled Electron Transfer versus Hydrogen Atom Transfer: Generation of Charge-Localized Diabatic States[‡]

2.1. Introduction

Proton-coupled electron transfer (PCET) reactions are ubiquitous throughout chemistry and biology.[1-5] Concerted PCET reactions involve the transfer of an electron and a proton in a single step without a stable intermediate. The relative timescales of the electrons, transferring proton, and solvent/protein environment dictate the appropriate theoretical treatment and form of the rate constant expression. Typically concerted PCET reactions are vibronically nonadiabatic because the quantum subsystem comprised of the electrons and transferring proton does not respond instantaneously to the solvent and protein motions. In this regime, PCET reactions are described in terms of nonadiabatic transitions between the reactant and product diabatic vibronic states, where the reactant (product) diabatic vibronic states correspond to the electron and proton localized on their donors (acceptors).[4,6]

The vibronic coupling is defined as the Hamiltonian matrix element between a pair of reactant and product diabatic vibronic states. In the vibronically nonadiabatic regime, the vibronic coupling is much less than the thermal energy.[4,7,8] The golden rule formalism has been

[‡]Reproduced with permission from:
A. Sirjoosingh, and S. Hammes-Schiffer, "Proton-Coupled Electron Transfer versus Hydrogen Atom Transfer: Generation of Charge-Localized Diabatic States," *J. Phys. Chem. A* **115**, 2367-2377 (2011). © 2011 American Chemical Society

used to derive nonadiabatic rate constant expressions, in which each term is proportional to the square of the relevant vibronic coupling.[9,10] Even within this regime, the proton transfer may be electronically adiabatic, where the electrons respond instantaneously to the proton, or electronically nonadiabatic, where the response of the electrons is slower than the proton motion. The form of the vibronic coupling differs in these two limits.[11,12]

Although hydrogen atom transfer (HAT) reactions may be viewed as a subset of concerted PCET reactions, often a distinction between HAT and PCET reactions is useful for the description of chemical reactions. Traditionally, HAT reactions are defined as the simultaneous transfer of an electron and proton between the same donor and acceptor without significant molecular charge redistribution, corresponding to small solvent reorganization energies. In contrast, PCET reactions typically involve different donors and acceptors for the electron and proton and are associated with significant molecular charge redistribution, corresponding to larger solvent reorganization energies. Distinguishing between the two types of mechanisms is important because they may require different rate constant expressions. In particular, due to the negligible solvent reorganization, HAT reactions may require an explicit dynamical treatment of intramolecular solute modes rather than the Marcus theory[13] description in terms of collective solvent coordinates typically applied to PCET reactions. In the literature, HAT and PCET mechanisms have been differentiated by the analysis of frontier molecular orbitals[14] and topographical characteristics of the potential energy surfaces.[15]

We have formulated a more quantitative diagnostic for distinguishing between HAT and PCET in terms of electronically adiabatic and nonadiabatic proton transfer, respectively.[12] The degree of electron-proton nonadiabaticity can be evaluated with a semiclassical formalism, where an adiabaticity parameter is defined as the ratio of the proton tunneling time and the

electronic transition time.[11] In the electronically adiabatic regime, the electronic transition time is much shorter than the proton tunneling time, so the proton transfer occurs on the electronically adiabatic ground state surface. In the electronically nonadiabatic regime, the proton tunneling time is much shorter than the electronic transition time, and excited electronic states are involved.

Our application of this semiclassical formalism to the self-exchange reactions in the phenoxyl-phenol and benzyl-toluene systems revealed that the former is electronically nonadiabatic, corresponding to PCET, while the latter is electronically adiabatic, corresponding to HAT.[12] These mechanistic insights were consistent with prior analyses of the frontier molecular orbitals for these systems.[14] Although the semiclassical approach is physically appealing, our previous implementation required fitting of the diabatic potential energy curves, which were represented by analytical functional forms, to electronic structure calculations of the adiabatic states. A more rigorous, well-defined method for constructing the diabatic states would be useful within this approach and, more generally, for the calculation of nonadiabatic PCET rate constants.

Numerous schemes have been developed for constructing diabatic states for electron transfer reactions. Mathematically, diabatic electronic states can be defined as states for which the derivative couplings vanish at all possible nuclear configurations.[16-38] Physically, the character of a diabatic state (i.e., its electronic charge density) does not change significantly as the nuclei move. For electron transfer, the two relevant diabatic states typically correspond to the electron localized on the donor and the acceptor, respectively (i.e., the reactant and product electronic states). A variety of approaches have been developed for constructing such charge-localized diabatic states for electron transfer reactions. These approaches include the

minimization of derivative couplings calculated in an adiabatic basis,[16,17] block diagonalization methods,[18,20,24] the generalized Mulliken-Hush method[23,24] and extensions with Boys localization,[32,34] constrained density functional theory (DFT),[30,31] and valence bond theory combined with molecular orbital or DFT methods.[35,36] To our knowledge, these electronic structure approaches for constructing diabatic states have not yet been applied to PCET reactions.

In this paper, we examine fundamental issues concerning the evaluation of electron-proton nonadiabaticity, the distinction between the HAT and PCET mechanisms, and the generation of charge-localized diabatic states for PCET reactions. In particular, we propose quantitative diagnostics of electron-proton nonadiabaticity that are directly accessible from standard quantum chemistry calculations. We also clarify the connection between the degree of electron-proton nonadiabaticity and the physical characteristics distinguishing HAT from PCET. Most importantly, we develop a rigorous theoretical approach to generate charge-localized reactant and product diabatic states for PCET reactions. In many PCET reactions, movement of the proton can induce the electron transfer reaction. Thus, we utilize the transferring hydrogen coordinate as a tool to construct charge-localized electronic states that are diabatic with respect to this hydrogen coordinate. Specifically, we generate the charge-localized diabatic states with an adiabatic-to-diabatic transformation matrix[16,17] defined to ensure that the first-order nonadiabatic coupling with respect to the transferring hydrogen coordinate vanishes.

An outline of the paper is as follows. Section 2.2A summarizes the theoretical formalism for calculating the electron-proton vibronic states in the double adiabatic representation. Section 2.2B presents a theoretical approach for generating electronic states that are rigorously diabatic with respect to the one-dimensional hydrogen coordinate, as well as the generation of the

associated reactant and product diabatic vibronic states relevant to PCET reactions. Section 2.2C describes the quantitative diagnostics of electron-proton nonadiabaticity that arise from the theoretical concepts of the first two subsections. The computational methods implemented to apply these theoretical approaches to the phenoxyl-phenol and benzyl-toluene systems are discussed in Section 2.3, and the results of these applications are presented in Section 2.4. Specifically, Section 2.4A describes the generation of the electronically adiabatic and diabatic potential energy curves and the calculation of the quantitative diagnostics for determining the degree of electron-proton nonadiabaticity. Section 2.4B presents an analysis of the adiabatic electronic states to clarify the connection between electron-proton nonadiabaticity and the physical characteristics of HAT and PCET. Section 2.4C presents an analysis of the diabatic electronic states for the PCET reaction to illustrate that these diabatic states exhibit physically meaningful, invariant charge localization along the transferring hydrogen coordinate. Concluding remarks are provided in Section 2.5.

2.2. Theory

A. Double adiabatic representation

Consider a system divided into N_e electrons, N_p protons, and N_s slow nuclei with coordinates $\mathbf{r}_e$, $\mathbf{r}_p$, and $\mathbf{R}$ and masses m_e, m_p and $\{M_I\}$, respectively, and with potential energy $V(\mathbf{r}_e, \mathbf{r}_p, \mathbf{R})$. The Hamiltonian for the "fast" degrees of freedom (i.e., the electron-proton subsystem) is

$$H_q = -\sum_{i'=1}^{N_p} \frac{\hbar^2}{2m_p} \nabla_{i'}^2 + H_e \,, \tag{2.1}$$

where the electronic Hamiltonian is

$$H_e = -\sum_{i=1}^{N_e} \frac{\hbar^2}{2m_e} \nabla_i^2 + V\left(\mathbf{r}_e, \mathbf{r}_p, \mathbf{R}\right). \tag{2.2}$$

For fixed $\mathbf{R}$, the eigenfunctions $\Phi_k\left(\mathbf{r}_e, \mathbf{r}_p; \mathbf{R}\right)$ of H_q are calculated by solving

$$H_q \Phi_k\left(\mathbf{r}_e, \mathbf{r}_p; \mathbf{R}\right) = E_k\left(\mathbf{R}\right) \Phi_k\left(\mathbf{r}_e, \mathbf{r}_p; \mathbf{R}\right). \tag{2.3}$$

Often this equation is analyzed in the context of a double adiabatic basis set. For this purpose, the adiabatic electronic states for fixed $\left(\mathbf{r}_p, \mathbf{R}\right)$ are determined by solving

$$H_e \psi_i\left(\mathbf{r}_e; \mathbf{r}_p, \mathbf{R}\right) = \varepsilon_i\left(\mathbf{r}_p, \mathbf{R}\right) \psi_i\left(\mathbf{r}_e; \mathbf{r}_p, \mathbf{R}\right). \tag{2.4}$$

The proton vibrational states for fixed $\mathbf{R}$ and adiabatic electronic state i are obtained by solving

$$\left(-\sum_{i'=1}^{N_p} \frac{\hbar^2}{2m_p} \nabla_{i'}^2 + \varepsilon_i\left(\mathbf{r}_p, \mathbf{R}\right) \right) \phi_\mu^{(i)}\left(\mathbf{r}_p; \mathbf{R}\right) = \varepsilon_\mu^{(i)}\left(\mathbf{R}\right) \phi_\mu^{(i)}\left(\mathbf{r}_p; \mathbf{R}\right), \tag{2.5}$$

where $\varepsilon_\mu^{(i)}\left(\mathbf{R}\right)$ is the energy of the double adiabatic electron-proton vibronic state $\left(i, \mu\right)$. The double adiabatic basis functions $\left\{\zeta_{i\mu}\left(\mathbf{r}_e, \mathbf{r}_p; \mathbf{R}\right)\right\}$ are defined as products of the adiabatic electronic and proton vibrational wavefunctions:

$$\zeta_{i\mu}\left(\mathbf{r}_e, \mathbf{r}_p; \mathbf{R}\right) = \psi_i\left(\mathbf{r}_e; \mathbf{r}_p, \mathbf{R}\right) \phi_\mu^{(i)}\left(\mathbf{r}_p; \mathbf{R}\right). \tag{2.6}$$

The double adiabatic approximation assumes the Born-Oppenheimer separation between the slow nuclei and the fast degrees of freedom and between the electrons and the protons. In other words, the electrons and protons are assumed to respond instantaneously to motions of the slow nuclei, and the electrons are assumed to respond instantaneously to motions of the protons. In this limit, the eigenfunctions of H_q are approximated as $\Phi_k\left(\mathbf{r}_e, \mathbf{r}_p; \mathbf{R}\right) \approx \zeta_{i\mu}\left(\mathbf{r}_e, \mathbf{r}_p; \mathbf{R}\right)$ with eigenvalues $E_k\left(\mathbf{R}\right) \approx \varepsilon_\mu^{(i)}\left(\mathbf{R}\right)$.

In PCET reactions, often the Born-Oppenheimer separation between electrons and quantum protons is not valid. In this case, Eq. (2.3) can be solved by expanding the eigenfunctions in the double adiabatic basis given by Eq. (2.6):

$$\Phi_k\left(\mathbf{r}_e,\mathbf{r}_p;\mathbf{R}\right)=\sum_{i,\mu}c_{i\mu}^k\zeta_{i\mu}\left(\mathbf{r}_e,\mathbf{r}_p;\mathbf{R}\right). \tag{2.7}$$

Following the standard linear variational procedure, we express the Hamiltonian H_q in the double adiabatic basis and diagonalize the resulting Hamiltonian matrix to obtain the eigenfunctions and eigenvalues of Eq. (2.3). The matrix elements in this basis are given by

$$\begin{aligned} H_{i\mu,j\nu} &= \left\langle \zeta_{i\mu} \mid H_q \mid \zeta_{j\nu} \right\rangle_{ep} \\ &= \delta_{ij}\delta_{\mu\nu}\varepsilon_\mu^{(i)}\left(\mathbf{R}\right) - \frac{\hbar^2}{m_p}\left\langle \phi_\mu^{(i)} \mid \mathbf{d}_{ij}^{(ep)}\cdot\nabla_{\mathbf{r}_p}\phi_\nu^{(j)} \right\rangle_p - \frac{\hbar^2}{2m_p}\left\langle \phi_\mu^{(i)} \mid g_{ij}^{(ep)} \mid \phi_\nu^{(j)} \right\rangle_p \end{aligned} \tag{2.8}$$

where

$$\mathbf{d}_{ij}^{(ep)}\left(\mathbf{r}_p;\mathbf{R}\right)=\left\langle \psi_i \mid \nabla_{\mathbf{r}_p}\psi_j \right\rangle_e = \begin{cases} \dfrac{\left\langle \psi_i \mid \nabla_{\mathbf{r}_p}H_e \mid \psi_j \right\rangle_e}{\varepsilon_j-\varepsilon_i} & i\neq j \\ 0 & i=j \end{cases} \tag{2.9}$$

is the first-order nonadiabatic coupling vector and

$$g_{ij}^{(ep)}\left(\mathbf{r}_p;\mathbf{R}\right)=\left\langle \psi_i \mid \nabla_{\mathbf{r}_p}^2\psi_j \right\rangle_e \tag{2.10}$$

is the second-order nonadiabatic coupling term. Here the notation $\left\langle\cdots\right\rangle_e$ and $\left\langle\cdots\right\rangle_p$ denotes integration over electronic coordinates $\mathbf{r}_e$ or proton coordinates $\mathbf{r}_p$, respectively. The Hamiltonian matrix expressed in the double adiabatic basis is diagonal only when the first- and second-order nonadiabatic coupling terms vanish. In this case, the Born-Oppenheimer separation between the electrons and protons is valid, and the proton motion is electronically adiabatic (i.e., the electrons respond instantaneously to motions of the protons).

Although the first-order nonadiabatic coupling terms can be calculated in many quantum chemistry programs, such as GAMESS,[39] the second-order nonadiabatic coupling terms are more problematic because they require second derivatives.[40-42] We have found that the second-order nonadiabatic couplings are often relatively large for PCET reactions and that neglecting these terms introduces significant asymmetry into the Hamiltonian matrix. Thus, we derived an alternative expression from Eq. (2.8) using the chain rule, integration by parts, and insertion of the identity operator:

$$H_{i\mu,j\nu} = \delta_{ij}\delta_{\mu\nu}\varepsilon_{\mu}^{(i)}(\mathbf{R}) - \frac{\hbar^2}{2m_p}\left[\left\langle \phi_{\mu}^{(i)} \mid \mathbf{d}_{ij}^{(ep)} \cdot \nabla_{\mathbf{r}_p} \phi_{\nu}^{(j)} \right\rangle_p + \left\langle \phi_{\nu}^{(j)} \mid \mathbf{d}_{ji}^{(ep)} \cdot \nabla_{\mathbf{r}_p} \phi_{\mu}^{(i)} \right\rangle_p\right] + \frac{\hbar^2}{2m_p}\sum_k \left\langle \phi_{\mu}^{(i)} \mid \mathbf{d}_{ki}^{(ep)} \cdot \mathbf{d}_{kj}^{(ep)} \mid \phi_{\nu}^{(j)} \right\rangle_p \tag{2.11}$$

where the summation in the last term is over all electronic states k. Assuming the electronic basis set is complete, this expression for the matrix elements is rigorously identical to Eq. (2.8) in that no terms have been neglected, but it avoids the calculation of second-order nonadiabatic coupling terms. Instead, this expression requires the calculation of first-derivative nonadiabatic coupling vectors between the states of interest, i and j, and all other electronic states, k. Typically we can assume that for $k > \max\{i, j\}$, the terms in this summation will be negligible compared to the other terms in Eq. (2.11) because the states k are much higher in energy. The main advantage of Eq. (2.11) is that it enables the calculation of vibronic state energies including both first-order and second-order nonadiabatic couplings without the explicit calculation of the second-derivative terms. A brief derivation of Eq. (2.11), as well as a comparison of implementations based on Eqs. (2.8) and (2.11), is provided in Appendix A.

B. Diabatic representation

As discussed in the Introduction, the theoretical description of PCET reactions is often based on the diabatic representation. In this subsection, we develop the methodology to obtain the relevant diabatic electronic states for PCET reactions. For N electronic states, we define $\vec{\psi}$ to be a column vector of the N electronic eigenfunctions $\left\{\psi_i\left(\mathbf{r}_e;\mathbf{r}_p,\mathbf{R}\right)\right\}$ of Eq. (2.4). Then we define an $N\times N$ transformation matrix $A\left(\mathbf{r}_p;\mathbf{R}\right)$ such that

$$\vec{\xi}\left(\mathbf{r}_e;\mathbf{r}_p,\mathbf{R}\right)=A\left(\mathbf{r}_p;\mathbf{R}\right)\vec{\psi}\left(\mathbf{r}_e;\mathbf{r}_p,\mathbf{R}\right), \tag{2.12}$$

where $\vec{\xi}$ is a column vector of functions satisfying the condition

$$\left\langle \xi_i \left| \nabla_{\mathbf{r}_p} \xi_j \right\rangle_e = 0 \quad \text{for all } i, j. \right. \tag{2.13}$$

Thus, the transformed electronic states $\left\{\xi_i\left(\mathbf{r}_e;\mathbf{r}_p,\mathbf{R}\right)\right\}$ satisfy the standard definition of diabatic states with respect to the proton coordinate $\mathbf{r}_p$. In Ref. 16, Baer derived the conditions on the matrix A for the general N-state case and showed that a transformation providing diabatic states that exactly satisfy Eq. (2.13) must satisfy the curl condition when the first-order nonadiabatic coupling vector is of dimension greater than unity.

For a one-dimensional proton coordinate r_p, however, we can form exact diabatic electronic states using the diabatization scheme outlined by Baer.[16] For simplicity we consider the case of $N=2$ electronic states, but the extension to more electronic states is possible.[29] As derived previously, the matrix A must satisfy

$$\frac{\partial A^t}{\partial r_p}+TA^t=0, \tag{2.14}$$

where the matrix T has elements

$$T_{ij}\left(r_p;\mathbf{R}\right)=d_{ij}^{(ep)}\left(r_p;\mathbf{R}\right). \tag{2.15}$$

Here, $d_{ij}^{(ep)}$ is given by Eq. (2.9), where $\nabla_{\mathbf{r}_p}$ is replaced by $\partial/\partial r_p$. Since Eq. (2.14) implies that A is orthogonal, we can express this matrix as

$$A\left(r_p;\mathbf{R}\right)=\begin{pmatrix}\cos\gamma & -\sin\gamma\\ \sin\gamma & \cos\gamma\end{pmatrix}. \tag{2.16}$$

To satisfy Eq. (2.14), $\gamma\left(r_p;\mathbf{R}\right)$ must satisfy

$$\frac{\partial\gamma}{\partial r_p}+d_{12}^{(ep)}=0. \tag{2.17}$$

The solution of Eq. (2.17) is of the form

$$\gamma\left(r_p;\mathbf{R}\right)=\gamma\left(r_0;\mathbf{R}\right)-\int_{r_0}^{r_p}d_{12}^{(ep)}\left(r;\mathbf{R}\right)\mathrm{d}r \tag{2.18}$$

where $\gamma\left(r_0;\mathbf{R}\right)$ is an additive constant that must be specified at some point $r_p=r_0$.

The diabatic potential energy matrix is given by $W=AUA^{-1}$, where $U_{ij}=\varepsilon_i\left(r_p,\mathbf{R}\right)\delta_{ij}$ is the adiabatic potential energy matrix. Substituting Eq. (2.16) into this expression for *W*, the diabatic potential energy matrix elements are expressed as

$$\begin{aligned}
W_{11}\left(r_p,\mathbf{R}\right)&=\varepsilon_1\left(r_p,\mathbf{R}\right)\cos^2\gamma\left(r_p;\mathbf{R}\right)+\varepsilon_2\left(r_p,\mathbf{R}\right)\sin^2\gamma\left(r_p;\mathbf{R}\right),\\
W_{22}\left(r_p,\mathbf{R}\right)&=\varepsilon_1\left(r_p,\mathbf{R}\right)\sin^2\gamma\left(r_p;\mathbf{R}\right)+\varepsilon_2\left(r_p,\mathbf{R}\right)\cos^2\gamma\left(r_p;\mathbf{R}\right),\\
W_{12}\left(r_p,\mathbf{R}\right)&=\left[\varepsilon_1\left(r_p,\mathbf{R}\right)-\varepsilon_2\left(r_p,\mathbf{R}\right)\right]\sin 2\gamma\left(r_p;\mathbf{R}\right)/2.
\end{aligned} \tag{2.19}$$

Here $W_{11}\left(r_p,\mathbf{R}\right)$ and $W_{22}\left(r_p,\mathbf{R}\right)$ are the diabatic electronic energies and $W_{12}\left(r_p,\mathbf{R}\right)$ is the diabatic electronic coupling.

Analogous to the double adiabatic vibronic basis set defined in Eq. (2.6), we define a diabatic vibronic basis set. The proton vibrational states for fixed $\mathbf{R}$ and diabatic electronic state i are obtained by solving

$$\left(-\frac{\hbar^2}{2m_p}\frac{\partial^2}{\partial r_p}+W_{ii}\left(r_p,\mathbf{R}\right)\right)\tilde{\varphi}_\mu^{(i)}\left(r_p;\mathbf{R}\right)=\tilde{\varepsilon}_\mu^{(i)}\left(\mathbf{R}\right)\tilde{\varphi}_\mu^{(i)}\left(r_p;\mathbf{R}\right), \tag{2.20}$$

where $\tilde{\varepsilon}_\mu^{(i)}\left(\mathbf{R}\right)$ is the energy of the diabatic electron-proton vibronic state $\left(i,\mu\right)$. The diabatic vibronic basis functions $\left\{\tilde{\zeta}_{i\mu}\left(\mathbf{r}_e,\mathbf{r}_p;\mathbf{R}\right)\right\}$ are defined as products of the diabatic electronic wavefunctions and associated proton vibrational wavefunctions:

$$\tilde{\zeta}_{i\mu}\left(\mathbf{r}_e,r_p;\mathbf{R}\right)=\xi_i\left(\mathbf{r}_e;r_p,\mathbf{R}\right)\tilde{\phi}_\mu^{(i)}\left(r_p;\mathbf{R}\right). \tag{2.21}$$

Eq. (2.3) can be solved by expanding the eigenfunctions in this diabatic vibronic basis. Expression of the Hamiltonian H_q in this basis leads to the matrix elements

$$\tilde{H}_{i\mu,j\nu}=\delta_{ij}\delta_{\mu\nu}\tilde{\varepsilon}_\mu^{(i)}\left(\mathbf{R}\right)+\left(1-\delta_{ij}\right)\left\langle\tilde{\varphi}_\mu^{(i)}\left|W_{ij}\right|\tilde{\varphi}_\nu^{(j)}\right\rangle_p, \tag{2.22}$$

and diagonalization of this Hamiltonian matrix provides the vibronic eigenfunctions and eigenvalues. This matrix diagonalization procedure leads to the same vibronic eigenfunctions and eigenvalues of Eq. (2.3) as those obtained using the double adiabatic basis.

C. Quantitative diagnostics of electron-proton nonadiabaticity

As discussed in the Introduction, PCET reactions may exhibit electronically adiabatic or nonadiabatic proton transfer, depending on the relative timescales of the electron and proton motions. In the electronically adiabatic regime, the electrons respond instantaneously to the proton motion and the system remains on the electronic ground state, while in the electronically nonadiabatic regime, the response of the electrons is slower than the proton motion and excited

electronic states are involved. The theoretical framework developed in the previous two subsections provides quantitative diagnostics of the electron-proton nonadiabaticity. For simplicity, we discuss the diagnostics in terms of only two electronic states (i.e., the ground and first excited adiabatic electronic states), although the extension to more electronic states is straightforward. We also consider only the two lowest-energy vibronic states and neglect mixing with higher-energy vibronic states. The extension to other pairs of vibronic states is also straightforward.

A useful quantity for characterizing PCET systems is the vibronic coupling, which was defined in the Introduction to be the Hamiltonian matrix element between a pair of reactant and product diabatic electron-proton vibronic states. For a symmetric system, the vibronic coupling between the ground reactant and product diabatic vibronic states is

$$V_{\mathrm{DA}} = \frac{E_2(\mathbf{R}) - E_1(\mathbf{R})}{2}, \tag{2.23}$$

where the two lowest-energy "exact" vibronic state energies E_1 and E_2 may be calculated by solution of Eq. (2.3) in either the double adiabatic or the diabatic vibronic basis.

In the double adiabatic representation, the degree of electron-proton nonadiabaticity is indicated by the magnitude of the off-diagonal terms in the Hamiltonian matrices defined equivalently by Eqs. (2.8) and (2.11). Another related diagnostic within the double adiabatic representation may also be useful. For electronically adiabatic proton transfer, the two lowest-energy "exact" vibronic states Φ_1 and Φ_2 possess predominantly ground electronic state character, while for electronically nonadiabatic proton transfer, these states possess more excited electronic state character. Thus, the degree of electronic nonadiabaticity for the proton transfer can be quantified by calculating the fraction of the excited electronic state character of each of

these vibronic states. The contribution of the adiabatic electronic state *i* to vibronic state *k* is quantified from the coefficients in Eq. (2.7) as

$$\theta_i\left(\Phi_k\right)=\sum_{\mu}\left(c_{i\mu}^{k}\right)^2 . \tag{2.24}$$

An additional quantitative diagnostic of electron-proton nonadiabaticity is provided by comparison of the "full" vibronic coupling given in Eq. (2.23) with the vibronic coupling in the electronically nonadiabatic and double adiabatic limits. In the electronically nonadiabatic limit, the vibronic coupling between the reactant and product diabatic vibronic states $\left(1,\mu\right)$ and $\left(2,\nu\right)$ is[11,12]

$$V_{\mathrm{DA}}^{(\mathrm{na})}=\left\langle \tilde{\varphi}_{\mu}^{(1)} \left| W_{12} \right| \tilde{\varphi}_{\nu}^{(2)} \right\rangle_{p} . \tag{2.25}$$

This expression reduces to the familiar form of the diabatic electronic coupling multiplied by the Franck-Condon overlap between the reactant and product proton vibrational wavefunctions when the diabatic electronic coupling $W_{12}\left(r_p,\mathbf{R}\right)$ defined in Eq. (2.19) is independent of r_p. In the double adiabatic limit, the vibronic coupling for a symmetric system is[11,12]

$$V_{\mathrm{DA}}^{(\mathrm{d\text{-}ad})}=\frac{\varepsilon_2^{(1)}\left(\mathbf{R}\right)-\varepsilon_1^{(1)}\left(\mathbf{R}\right)}{2} , \tag{2.26}$$

where $\varepsilon_{\mu}^{(i)}\left(\mathbf{R}\right)$ are the energies of the double adiabatic electron-proton vibronic states defined by Eq. (2.6). Thus, the double adiabatic vibronic coupling corresponds to the energy splitting between the lowest two proton vibrational states on the ground adiabatic electronic state.

2.3. Computational Methods

We used the phenoxyl-phenol and benzyl-toluene self-exchange reactions as model systems to illustrate the theoretical concepts developed above. The transition state geometries

for these systems were obtained from Ref. 14 using density functional theory (DFT) with the B3LYP functional[43,44] and the 6-31G* basis set.[45] In the present calculations, all nuclei except the transferring hydrogen atom were fixed at the transition state geometry. For each system, we calculated the two lowest-energy electronically adiabatic potential energy curves for hydrogen atom positions on a grid spanning the hydrogen donor-acceptor axis using the complete active space self-consistent field (CASSCF) method. The CASSCF calculations were performed using the 6-31G* basis set and state-averaging over the ground and first excited electronic states with equal weighting. An active space of three electrons in six orbitals was chosen at each grid point to maintain the character of the orbitals along the hydrogen coordinate. Note that these calculations were performed at a relatively low level of theory because our objective was to utilize these systems as simple models to illustrate the qualitative features of the electronic wavefunctions and potential energy curves. We obtained the nonadiabatic coupling vectors directly from the CASSCF calculations. All of the CASSCF calculations were performed using the GAMESS electronic structure package.[39]

We calculated the proton vibrational wavefunctions corresponding to both the adiabatic and diabatic electronic potential energy surfaces by solving the one-dimensional Schrödinger equations given in Eq. (2.5) and Eq. (2.20), respectively. These calculations were performed with the Fourier grid Hamiltonian method[46,47] using 128 grid points along the hydrogen donor-acceptor axis. For the construction of the vibronic Hamiltonians in the double adiabatic and diabatic vibronic bases given in Eqs. (2.11) and (2.22), respectively, forty proton wavefunctions for each electronic state were included.

To aid in the analysis of charge transfer properties, we calculated the dipole moments, atomic charges, and electrostatic potential maps as functions of the hydrogen coordinate for the

adiabatic and diabatic electronic states. The properties of the adiabatic electronic states were calculated directly from the CASSCF wavefunctions with GAMESS. For the diabatic electronic states, we modified a local version of GAMESS to calculate these properties for the appropriate linear combination of configuration interaction (CI) states following the transformation given in Eq. (2.12). The atomic charges were obtained by fitting to the electrostatic potential calculated at points on the Connolly surface[48] under the constraint of reproducing the total charge and dipole moment of the electronic state under consideration.[49]

2.4. Results

A. Electron-proton nonadiabaticity

Figure 2.1 depicts the electronically adiabatic and diabatic potential energy curves as functions of the hydrogen coordinate for the phenoxyl-phenol and benzyl-toluene systems. The solid black curves are the ground and first excited state adiabatic electronic energies $\varepsilon_1\left(r_p,\mathbf{R}\right)$ and $\varepsilon_2\left(r_p,\mathbf{R}\right)$, respectively, calculated with the CASSCF method. The blue and red dashed curves are the diabatic electronic energies $W_{11}\left(r_p,\mathbf{R}\right)$ and $W_{22}\left(r_p,\mathbf{R}\right)$, respectively, calculated from the expressions in Eq. (2.19). The electronically adiabatic curves are qualitatively similar to those presented in Ref. 12 but differ slightly due to use of the 6-31G* instead of the 6-31G basis set. The diabatic states in the previous work were obtained by fitting the electronically adiabatic curves with specific functional forms in an empirical valence bond potential.[12] In the present work, however, the diabatic states were generated from the mathematical formulation given in Section 2.2B, thereby avoiding the fitting procedure and the assumption of a particular functional form.

In particular, the diabatic electronic wavefunctions were calculated with Eq. (2.12), and the corresponding diabatic electronic energies and coupling were calculated with Eq. (2.19) using the electronically adiabatic potential energy curves and the first-order nonadiabatic coupling vectors. This procedure ensures that the diabatic states rigorously satisfy the diabaticity condition given in Eq. (2.13) along the one-dimensional hydrogen coordinate. To obtain physically meaningful diabatic states, we chose $r_0 = 0$, corresponding to the transition state geometry, and set $\gamma(r_0) = -\pi/4$ in Eq. (2.18). This choice ensures that the adiabatic electronic states mix maximally at the transition state geometry and that the magnitude of the diabatic electronic coupling, W_{12}, is exactly half the splitting between the adiabatic electronic energies at this geometry. Note that the diabatic electronic potential energy W_{11} has a minimum corresponding to the transferring hydrogen localized on the donor, while the diabatic electronic potential energy W_{22} has a minimum corresponding to the transferring hydrogen localized on the acceptor. Thus, these two diabatic states correspond to the physically meaningful reactant and product states in a PCET reaction.

Table 2.1 presents the vibronic couplings calculated with a range of methods for the phenoxyl-phenol and benzyl-toluene systems. The full vibronic coupling, V_{DA}, was calculated with Eq. (2.23), where the exact vibronic state energies E_k were calculated by solution of Eq. (2.3) through full basis set diagonalization of the Hamiltonian matrix. The value of V_{DA} was found to be the same to within less than 0.1 cm^{-1} using either the double adiabatic vibronic basis (i.e., Hamiltonian matrix elements given by Eq. (2.11)) or the diabatic vibronic basis (i.e., Hamiltonian matrix elements given by Eq. (2.22)). In forming the double adiabatic vibronic Hamiltonian, we neglected the last term in Eq. (2.11) for $i \neq j$ because all terms involving $d_{k1}^{(ep)}$

and $d_{k2}^{(ep)}$ vectors for $k > 2$ were found to be much smaller than $d_{12}^{(ep)}$. More information about the double adiabatic vibronic Hamiltonian implementation is provided in Appendix A.

For comparison, we calculated the double adiabatic vibronic coupling, $V_{\mathrm{DA}}^{(\mathrm{d\text{-}ad})}$, using Eq. (2.26) and the nonadiabatic vibronic coupling, $V_{\mathrm{DA}}^{(\mathrm{na})}$, using Eq. (2.25). We also calculated the semiclassical vibronic coupling, $V_{\mathrm{DA}}^{(\mathrm{sc})}$, using the methodology presented in Ref. 11 with parameters obtained from the exact diabatic electronic energies and couplings derived above. For both systems, the full vibronic coupling agrees well with the semiclassical vibronic coupling. For the phenoxyl-phenol system, the nonadiabatic vibronic coupling agrees well with the full vibronic coupling, and for the benzyl-toluene system, the double adiabatic vibronic coupling agrees well with the full vibronic coupling. These calculations confirm that the phenoxyl-phenol system is electronically nonadiabatic, while the benzyl-toluene system is electronically adiabatic. Thus, a comparison of the nonadiabatic and double adiabatic vibronic couplings to the full vibronic coupling provides a diagnostic of the degree of electron-proton nonadabaticity, depending on which limiting expression agrees with the full vibronic coupling.

Another quantitative measure of the degree of electron-proton nonadiabaticity is the magnitude of the nonadiabatic coupling terms in Eq. (2.11). Figure 2.2 depicts the component of the first-order nonadiabatic coupling vector along the hydrogen donor-acceptor axis for the phenoxyl-phenol and benzyl-toluene systems. The magnitude of the first-order nonadiabatic coupling is significantly greater for the phenoxyl-phenol system, supporting the identification of the phenoxyl-phenol and benzyl-toluene mechanisms as electronically nonadiabatic and adiabatic, respectively.

Furthermore, the degree of electron-proton nonadiabaticity can be quantified by the fraction of excited adiabatic electronic state character of each vibronic state, as defined in Eq.

(2.24). The fractional contributions of the first excited adiabatic electronic state (i=2) to the lowest two vibronic states (k = 1 and 2) are given in Table 2.1. Note that the contributions of the first excited electronic state are significantly greater for the phenoxyl-phenol system than for the benzyl-toluene system. These results are consistent with the benzyl-toluene reaction occurring on the ground adiabatic electronic state and the phenoxyl-phenol reaction involving excited electronic states.

B. Distinction between PCET and HAT mechanisms

To obtain further insight into the distinction between the PCET and HAT mechanisms, we examined the charge transfer properties for the ground adiabatic electronic states of the phenoxyl-phenol and benzyl-toluene systems. For this analysis, the donor (acceptor) is the molecule that is protonated when the hydrogen coordinate has negative (positive) values. The reactant corresponds to the protonated donor, and the product corresponds to the protonated acceptor (i.e., the two minima in Figure 2.1). Figure 2.3 depicts the component of the dipole moment along the hydrogen donor-acceptor axis as the hydrogen is transferred. For both systems, the sign of this component of the dipole moment vector changes over the course of the reaction. The magnitude of the change in the dipole moment during the reaction, however, is significantly larger for the phenoxyl-phenol system than for the benzyl-toluene system. This figure indicates a much greater change in the electronic charge distribution for the phenoxyl-phenol system.

Figure 2.4 depicts the charges on the donor and acceptor molecules, as well as the transferring hydrogen, as the reaction proceeds. The charges on the donor and acceptor molecules switch signs during the reaction for the phenoxyl-phenol system but not for the

benzyl-toluene system. Moreover, the changes in the charges on the donor and acceptor molecules during the reaction are much greater for the phenoxyl-phenol system than for the benzyl toluene system. In addition, the charge on the transferring hydrogen is more positive for the phenoxyl-phenol system (~0.6 *e*) than for the benzyl-toluene system (~0.25 *e*). This figure illustrates that the electronic charge localizes on the donor or acceptor molecule for the phenoxyl-phenol system but not for the benzyl-toluene system. This figure also indicates that the transferring hydrogen has more positively charged proton character for the phenoxyl-phenol system than for the benzyl-toluene system. This analysis is consistent with the PCET mechanism for the phenoxyl-phenol system and the HAT mechanism for the benzyl-toluene system. For the PCET mechanism, a proton (i.e., a positively charged hydrogen) transfers simultaneously with electron transfer from the donor to the acceptor molecule. For the HAT mechanism, a hydrogen atom (i.e., a nearly neutral hydrogen) transfers without significantly changing the electronic charge on the donor and acceptor molecules.

The trends observed in Figures 2.3 and 2.4 are confirmed by the electrostatic potential maps depicted in Figure 2.5 for the reactant, transition state, and product positions of the transferring hydrogen. For the phenoxyl-phenol system, excess electronic charge is localized on the donor (acceptor) molecule when the proton is near the donor (acceptor). As the proton transfers from the donor to the acceptor, the electronic charge shifts from the donor to the acceptor molecule. This observation is consistent with the substantial changes in the dipole moment and the donor/acceptor charges shown in Figures 2.3 and 2.4. In contrast, for the benzyl-toluene system, the electronic charge distribution is more evenly distributed between the donor and acceptor molecules during the entire hydrogen transfer reaction. This observation is consistent with the relatively small changes in the dipole moment and the donor/acceptor charges

shown in Figures 2.3 and 2.4. These electrostatic potential maps highlight the PCET mechanism of the phenoxyl-phenol system and contrast it with the HAT mechanism of the benzyl-toluene system.

The change in the electronic charge distribution during the charge transfer reaction is related to the electron-proton nonadiabaticity discussed in the previous subsection. As indicated by Eq. (2.9), the nonadiabatic coupling terms within the double adiabatic representation reflect the change in the character of the adiabatic electronic wavefunction along the transferring hydrogen coordinate. Thus, significant electronic charge redistribution during hydrogen transfer is associated with a large nonadiabatic coupling, which in turn is a diagnostic for substantial electron-proton nonadiabaticity. Since PCET reactions are associated with greater charge redistribution than HAT reactions, this analysis clarifies the connection between the degree of electron-proton nonadiabaticity and the physical characteristics distinguishing PCET from HAT.

C. Properties of diabatic states

To illustrate the physical aspects of the diabatization scheme presented in Section 2.2B, we examined the charge transfer properties of the diabatic electronic states and compared them to those examined in the previous subsection for the ground adiabatic electronic state. We focus on the phenoxyl-phenol system because the diabatic representation is more applicable to the PCET mechanism than to the HAT mechanism. The charge transfer properties of the diabatic electronic states for the benzyl-toluene system are provided in Appendix A and are qualitatively very similar to the corresponding properties of the ground adiabatic electronic state.

The charge transfer properties of the diabatic electronic states are illustrated in Figures 2.6-2.8. Figure 2.6 depicts the component of the dipole moment along the hydrogen donor-

acceptor axis for the ground and excited adiabatic electronic states (solid black curves) and the two diabatic electronic states (blue and red dashed curves). This component of the dipole moment is positive for diabatic state ξ_1 and negative for diabatic state ξ_2 at all positions of the transferring hydrogen (i.e., during the entire reaction). Similarly, Figure 2.7 illustrates that the charge on the donor molecule is negative and the charge on the acceptor molecule is positive along the entire range of hydrogen positions for the reactant diabatic state ξ_1. The analogous figure for diabatic state ξ_2, where the charge on the donor molecule remains positive and the charge on the acceptor molecule remains negative during the entire reaction, is provided in Appendix A. Similar trends are exhibited in Figure 2.8, where the electrostatic potential maps appear virtually identical for the reactant, transition state, and product positions of the hydrogen for each diabatic state ξ_1 and ξ_2.

These figures for the diabatic electronic states are in sharp contrast to the analogous figures for the adiabatic electronic states discussed in the previous subsection. In particular, the donor and acceptor charges depicted in Figure 2.4a and the electrostatic maps depicted in Figure 2.5a for the adiabatic electronic states show substantial changes in the electronic charge distribution during the PCET reaction. Thus, the diabatization scheme successfully transforms the adiabatic electronic states into diabatic electronic states with relatively invariant charge distributions along the hydrogen coordinate. This property of charge invariance is important when using these diabatic states in PCET theories based on the nonadiabatic golden rule formalism.[4]

The diabatization scheme used in these studies may be extended in several directions. For the systems studied in the present paper, the proton coordinate was defined along the axis connecting the proton donor and acceptor. In general, the one-dimensional coordinate along

which the nonadiabatic coupling vanishes could be chosen to be a curved path and could involve motions of other nuclei in addition to the proton. For example, this coordinate could be chosen to be a normal mode coordinate dominated by the proton motion. In these cases, the implementation of the diabatization scheme would require expansion of the chosen coordinate as a linear combination of Cartesian coordinates or numerical calculation of the nonadiabatic coupling along the chosen coordinate. Furthermore, for these symmetric systems, we chose $r_0 = 0$, corresponding to the transition state geometry, and set $\gamma(r_0) = -\pi/4$. Note that the transformed states will be rigorously diabatic (i.e., will have zero first-derivative couplings) for any choice of $\gamma(r_0)$, but our objective is to choose a value that provides physically meaningful, charge-localized diabatic states. For asymmetric systems, we could choose r_0 to be the coordinate corresponding to the maximum of d_{12} and set $\gamma(r_0) = -\pi/4$ to ensure that the adiabatic states mix maximally and the diabatic states cross at the coordinate where the nonadiabatic coupling is the largest. All of these directions are interesting topics for future research.

2.5. Conclusions

In this paper, we devised quantitative diagnostics to evaluate the degree of electron-proton nonadiabaticity in PCET systems. One diagnostic is the comparison of the vibronic coupling calculated with the full basis set diagonalization method to the vibronic couplings calculated with the double adiabatic and nonadiabatic methods. Another diagnostic is the investigation of the magnitude of the nonadiabatic coupling terms within the double adiabatic representation. A third diagnostic is the calculation of the fraction of excited electronic state character in the relevant vibronic states. Application of these diagnostics to the phenoxyl-phenol

and benzyl-toluene systems confirmed that the former corresponds to electronically nonadiabatic and the latter corresponds to electronically adiabatic proton transfer.

Our calculations also clarified the connection between the degree of electron-proton nonadiabaticity and the charge transfer characteristics distinguishing PCET from HAT. According to traditional definitions, the extent of electronic charge redistribution is significantly greater for PCET than for HAT. Thus, analysis of the dipole moment, partial atomic charges, and electrostatic potential maps for the ground state adiabatic electronic wavefunctions along the transferring hydrogen coordinate can be used to differentiate the PCET and HAT mechanisms. Comparison of these charge transfer properties for the phenoxyl-phenol and benzyl-toluene systems clearly designated the former as PCET and the latter as HAT. Furthermore, the significant change in charge transfer character of the ground and first excited state electronic wavefunctions along the transferring hydrogen coordinate is associated with a large nonadiabatic coupling between these two electronic states. In turn, this large nonadiabatic coupling is a diagnostic associated with significant electron-proton nonadiabaticity. Thus, the extent of electronic charge redistribution during the reaction is related to the degree of electron-proton nonadiabaticity, and both properties may be utilized to differentiate PCET from HAT mechanisms. Specifically, PCET and HAT mechanisms correspond to electronically nonadiabatic and adiabatic proton transfer, respectively. As a result, the nonadiabatic expression for the vibronic coupling is applicable to PCET reactions, while the double adiabatic expression for the vibronic coupling is applicable to HAT reactions.

In addition, we developed a rigorous diabatization scheme for transforming the adiabatic electronic states generated from standard quantum chemistry calculations into charge-localized diabatic states for PCET reactions. These diabatic states are constructed so that the first-order

nonadiabatic couplings with respect to the one-dimensional transferring hydrogen coordinate vanish exactly. The application of this scheme to the phenoxyl-phenol system illustrated that the resulting diabatic states possess physically meaningful, localized electronic charge distributions that are relatively invariant along the hydrogen coordinate. These diabatic electronic states can be combined with the associated proton vibrational states to generate the reactant and product electron-proton vibronic states that form the basis of nonadiabatic PCET theories. Furthermore, these vibronic states and the corresponding vibronic couplings may be used to calculate rate constants and kinetic isotope effects of PCET reactions for comparison to experimental measurements.

System	V_{DA}	$V_{DA}^{(d\text{-}ad)}$	$V_{DA}^{(na)}$	$V_{DA}^{(sc)}$	$\theta_2(\Phi_1)$	$\theta_2(\Phi_2)$
Phenol	7.3	27.6	7.2	7.3	$1.13\text{x}10^{-4}$	$2.21\text{x}10^{-3}$
Toluene	20.9	21.0	37.6	20.6	$3.44\text{x}10^{-6}$	$6.69\text{x}10^{-6}$

Table 2.1: Vibronic couplings in cm^{-1} for the phenoxyl-phenol and benzyl-toluene systems calculated with various methods: Eq. (2.23) after full basis set diagonalization to obtain V_{DA}, Eq. (2.26) based on the double adiabatic approximation to obtain $V_{DA}^{(d\text{-}ad)}$, Eq. (2.25) in the electronically nonadiabatic limit to obtain $V_{DA}^{(na)}$, and the semiclassical formalism to obtain $V_{DA}^{(sc)}$. The last two columns provide the fractional contribution from the first excited adiabatic electronic state (i=2) to the lowest two vibronic states (k = 1 and 2), as defined in Eq. (2.24).

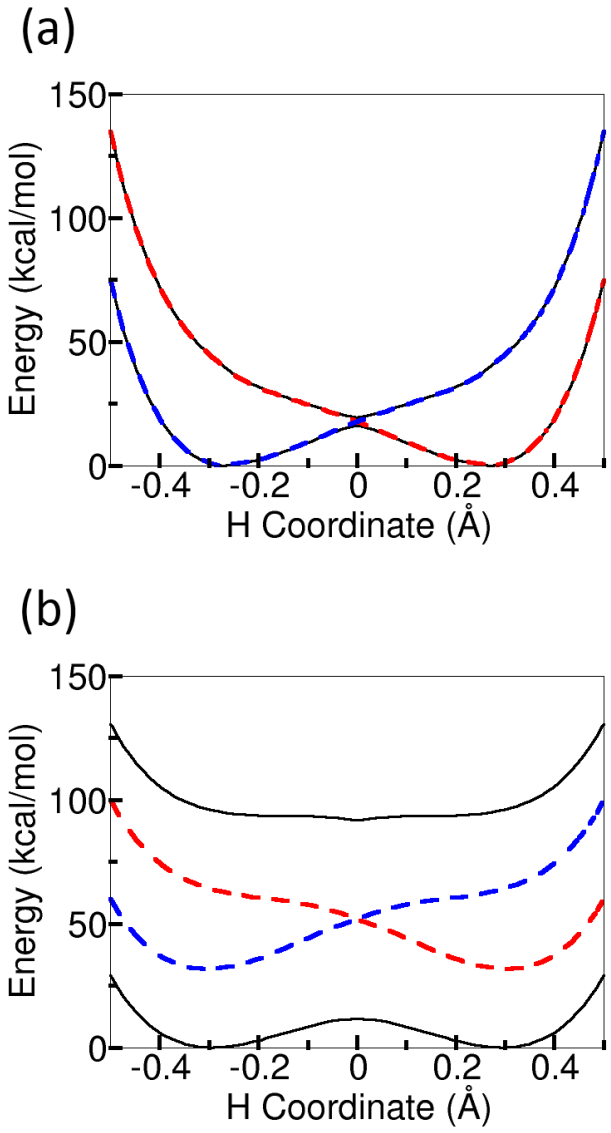

Figure 2.1: Electronically adiabatic and diabatic potential energy curves as functions of the hydrogen coordinate for the (a) phenoxyl-phenol and (b) benzyl-toluene systems. The solid black curves are the ground and first excited state adiabatic energies $\varepsilon_1\left(r_p,\mathbf{R}\right)$ and $\varepsilon_2\left(r_p,\mathbf{R}\right)$, respectively, calculated with the CASSCF method. The dashed blue and red curves are the diabatic electronic energies $W_{11}\left(r_p,\mathbf{R}\right)$ and $W_{22}\left(r_p,\mathbf{R}\right)$, respectively, calculated from the expressions in Eq. (2.19) with $\gamma\left(r_0=0\right)=-\pi/4$.

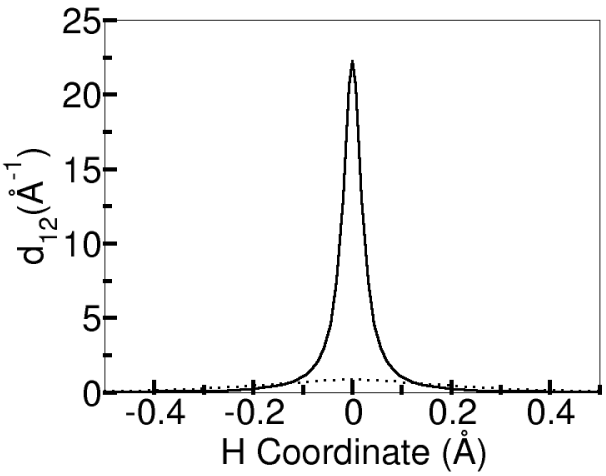

Figure 2.2: Component of the first-order nonadiabatic coupling vector, as defined in Eq. (2.9), between the ground and first excited adiabatic electronic states along the hydrogen donor-acceptor axis for the phenoxyl-phenol (solid) and benzyl-toluene (dotted) systems.

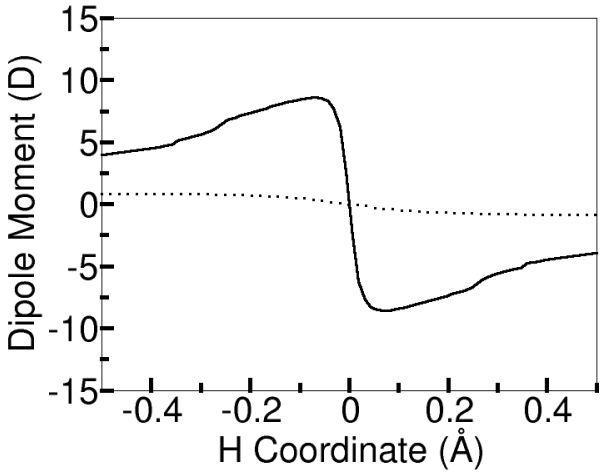

Figure 2.3: Component of the dipole moment vector along the hydrogen donor-acceptor axis for the ground adiabatic electronic state of the phenoxyl-phenol (solid) and benzyl-toluene (dotted) systems. A positive (negative) dipole moment indicates a dipole moment vector pointing toward the acceptor (donor).

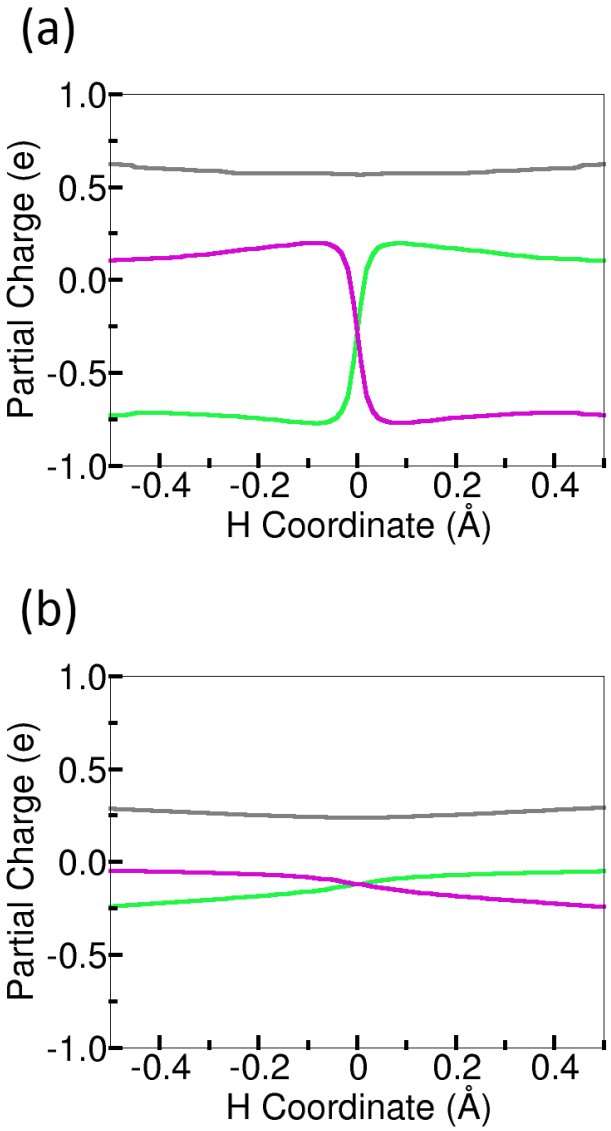

Figure 2.4: Partial charges determined from electrostatic potential-derived atomic charges for the ground adiabatic electronic state of the (a) phenoxyl-phenol and (b) benzyl-toluene systems. Partial charges are shown for the donor molecule (green), acceptor molecule (purple), and transferring hydrogen (grey).

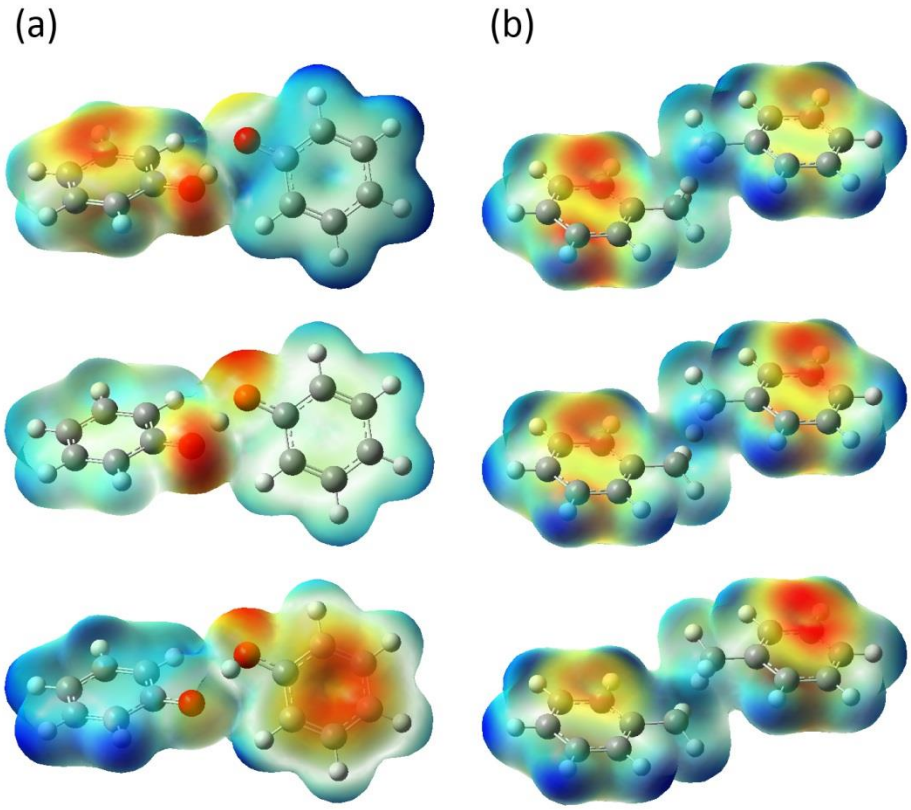

Figure 2.5: Electrostatic potential maps for the ground adiabatic electronic states corresponding to a density isosurface value of 0.005 for the reactant (top), transition state (middle), and product (bottom) positions of the transferring hydrogen for the (a) phenoxyl-phenol and (b) benzyl-toluene systems. Negatively and positively charged regions are indicated by red and blue coloring, respectively.

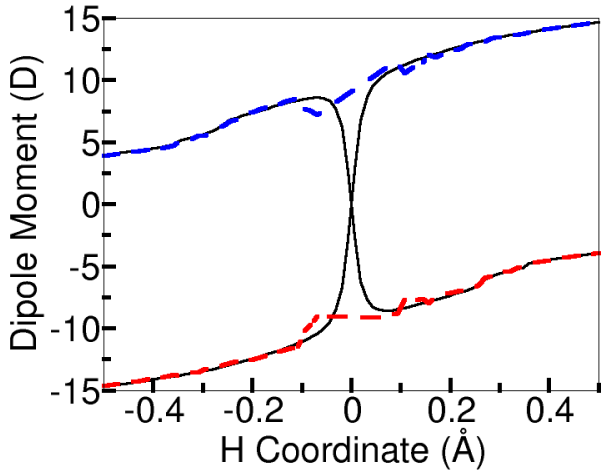

Figure 2.6: Component of the dipole moment vector along the hydrogen donor-acceptor axis for the ground and first-excited adiabatic electronic states (solid black curves) and the diabatic electronic states W_{11} (dashed blue curve) and W_{22} (dashed red curve) for the phenoxyl-phenol system. Calculated values of the dipole moment for the diabatic electronic states around $r_p = 0$ are omitted due to numerical noise in this region. A positive (negative) dipole moment indicates a dipole moment vector pointing toward the acceptor (donor).

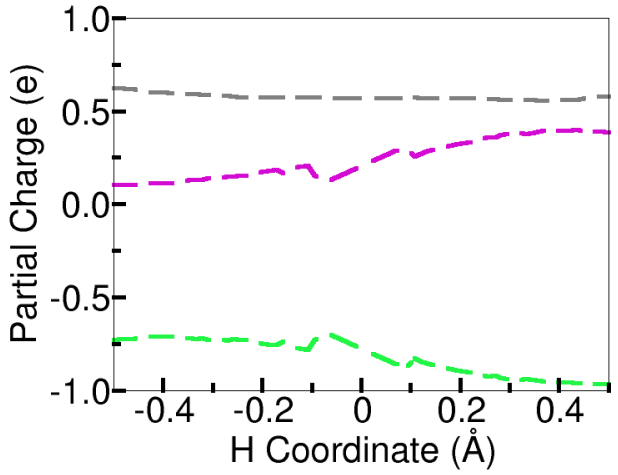

Figure 2.7: Partial charges determined from electrostatic potential-derived atomic charges for the diabatic electronic state W_{11} of the phenoxyl-phenol system. Partial charges are shown for the donor molecule (green), acceptor molecule (purple), and transferring hydrogen (grey). Calculated values of the partial charges around $r_p = 0$ are omitted due to numerical noise in this region. The corresponding plot for the diabatic electronic state W_{22} is essentially the reflection of this plot about the vertical axis corresponding to $r_p = 0$ with the donor and acceptor partial charges interchanged (i.e., the green and purple curves interchanged).

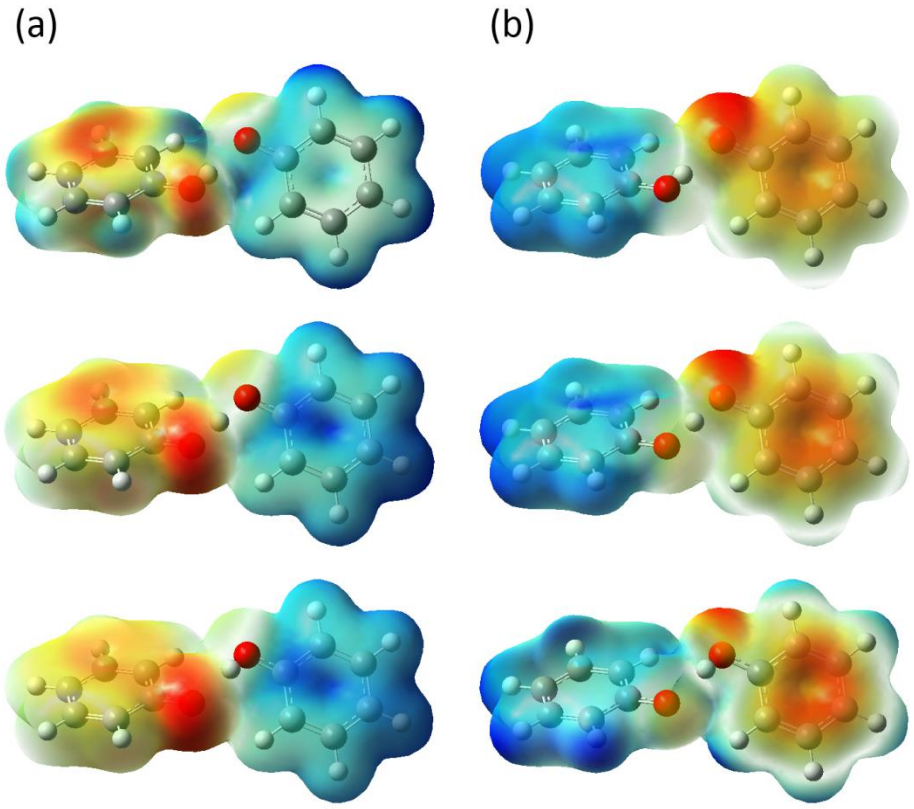

Figure 2.8: Electrostatic potential maps for the diabatic electronic states (a) W_{11} and (b) W_{22} corresponding to a density isosurface value of 0.005 for the reactant (top), transition state (middle), and product (bottom) positions of the transferring hydrogen for the phenoxyl-phenol system. Negatively and positively charged regions are indicated by red and blue coloring, respectively.

References

(1) R. I. Cukier and D. G. Nocera, *Annual Review of Physical Chemistry* **49**, 337 (1998).
(2) J. M. Mayer, *Annual Review of Physical Chemistry* **55**, 363 (2004).
(3) M. H. Huynh and T. J. Meyer, *Chemical Reviews* **107**, 5004 (2007).
(4) S. Hammes-Schiffer and A. V. Soudackov, *Journal of Physical Chemistry B* **112**, 14108 (2008).
(5) S. Hammes-Schiffer and A. A. Stuchebrukhov, *Chemical Reviews* **110**, 6939 (2010).
(6) S. Hammes-Schiffer, *Accounts of Chemical Research* **34**, 273 (2001).
(7) A. V. Barzykin, P. A. Frantsuzov, K. Seki, and M. Tachiya, *Advances in Chemical Physics* **123**, 511 (2002).
(8) A more general definition of nonadiabatic is that the characteristic time of passing through the nonadiabatic region is short compared to the characteristic time of nonadiabatic transitions. The golden rule formalism is valid when the vibronic coupling is much smaller than the thermal energy and the solvent relaxation is sufficiently fast.
(9) A. Soudackov and S. Hammes-Schiffer, *Journal of Chemical Physics* **113**, 2385 (2000).
(10) A. Soudackov, E. Hatcher, and S. Hammes-Schiffer, *Journal of Chemical Physics* **122**, 014505 (2005).
(11) Y. Georgievskii and A. A. Stuchebrukhov, *Journal of Chemical Physics* **113**, 10438 (2000).
(12) J. H. Skone, A. V. Soudackov, and S. Hammes-Schiffer, *Journal of the American Chemical Society* **128**, 16655 (2006).
(13) R. A. Marcus and N. Sutin, *Biochimica et Biophysica Acta* **811**, 265 (1985).
(14) J. M. Mayer, D. A. Hrovat, J. L. Thomas, and W. T. Borden, *Journal of the American Chemical Society* **124**, 11142 (2002).
(15) O. Tishchenko, D. G. Truhlar, A. Ceulemans, and M. T. Nguyen, *Journal of the American Chemical Society* **130**, 7000 (2008).
(16) M. Baer, *Chemical Physics Letters* **35**, 112 (1975).
(17) M. Baer, *Molecular Physics* **40**, 1011 (1980).
(18) T. Pacher, L. S. Cederbaum, and H. Koppel, *Journal of Chemical Physics* **89**, 7367 (1988).
(19) M. D. Newton, *Chemical Reviews* **91**, 767 (1991).
(20) T. Pacher, L. S. Cederbaum, and H. Koppel, *Advances in Chemical Physics* **84**, 293 (1993).
(21) K. Ruedenberg and G. J. Atchity, *Journal of Chemical Physics* **99**, 3799 (1993).
(22) T. A. Wesolowski and A. Warshel, *Journal of Physical Chemistry* **97**, 8050 (1993).
(23) R. J. Cave and M. D. Newton, *Chemical Physics Letters* **249**, 15 (1996).
(24) R. J. Cave and M. D. Newton, *Journal of Chemical Physics* **106**, 9213 (1997).
(25) G. J. Atchity and K. Ruedenberg, *Theoretical Chemistry Accounts* **97**, 47 (1997).
(26) W. Thiel and H. Koppel, *Journal of Chemical Physics* **110**, 9371 (1999).
(27) H. Nakamura and D. G. Truhlar, *Journal of Chemical Physics* **115**, 10353 (2001).
(28) H. Nakamura and D. G. Truhlar, *Journal of Chemical Physics* **117**, 5576 (2002).
(29) M. Baer, *Physics Reports* **358**, 75 (2002).
(30) Q. Wu and T. Van Voorhis, *Journal of Chemical Physics* **125**, 164105 (2006).
(31) Q. Wu and T. Van Voorhis, *Journal of Chemical Theory and Computation* **2**, 765 (2006).

(32) J. E. Subotnik, S. Yeganeh, R. J. Cave, and M. A. Ratner, *Journal of Chemical Physics* **129**, 244101 (2008).
(33) B. N. Papas, M. S. Schuurman, and D. R. Yarkony, *Journal of Chemical Physics* **129**, 124104 (2008).
(34) J. E. Subotnik, R. J. Cave, R. P. Steele, and N. Shenvi, *Journal of Chemical Physics* **130**, 234102 (2009).
(35) L. Song, Y. Mo, and J. Gao, *Journal of Chemical Theory and Computation* **5**, 174 (2009).
(36) A. Cembran, L. Song, Y. Mo, and J. Gao, *Journal of Chemical Theory and Computation* **5**, 2702 (2009).
(37) T. van Voorhis, T. Kowalczyk, B. Kaduk, L.-P. Wang, C.-L. Cheng, and Q. Wu, *Annual Reviews of Physical Chemistry* **61**, 149 (2010).
(38) X. Zhu and D. R. Yarkony, *Journal of Chemical Physics* **132**, 104101 (2010).
(39) M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, *Journal of Computational Chemistry* **14**, 1347 (1993).
(40) J. Morelli and S. Hammes-Schiffer, *Chemical Physics Letters* **269**, 161 (1997).
(41) I. Lengsfield, B. H. and D. R. Yarkony, *Journal of Chemical Physics* **84**, 348 (1986).
(42) E. F. Valeev and C. D. Sherrill, *Journal of Chemical Physics* **118**, 3921 (2003).
(43) A. D. Becke, *Journal of Chemical Physics* **98**, 5648 (1993).
(44) C. Lee, W. Yang, and P. G. Parr, *Physical Review B* **37**, 785 (1988).
(45) P. C. Hariharan and J. A. Pople, *Theoretica Chimica Acta* **28**, 213 (1973).
(46) C. C. Marston and G. G. Balint-Kurti, *Journal of Chemical Physics* **91**, 3571 (1989).
(47) S. P. Webb and S. Hammes-Schiffer, *Journal of Chemical Physics* **113**, 5214 (2000).
(48) M. L. Connolly, *Journal of Applied Crystallography* **16**, 548 (1983).
(49) U. C. Singh and P. A. Kollman, *Journal of Computational Chemistry* **5**, 129 (1984).

Chapter 3

Diabatization Schemes for Generating Charge-Localized Electron-Proton Vibronic States in Proton-Coupled Electron Transfer Systems[‡]

3.1. Introduction

Proton-coupled electron transfer (PCET) reactions are prevalent in many facets of biology, chemistry and physics.[1-5] Concerted PCET reactions, which are characterized by simultaneous electron and proton transfer without a stable intermediate, have been observed in a wide variety of systems, including enzymatic, photoinduced, and electrochemical processes. Typically PCET reactions are characterized as vibronically nonadiabatic because the subsystem comprised of the electrons and transferring proton does not respond instantaneously to the motions of the solvent and other solute nuclei. Within this nonadiabatic framework, PCET theories require the identification of charge-localized reactant and product diabatic electron-proton vibronic states corresponding to the transferring electron and proton localized on their donors and acceptors, respectively.[2,6] These diabatic electron-proton vibronic states can be used to calculate the vibronic couplings, which are key quantities in the nonadiabatic PCET rate constant expressions derived with the golden rule formalism.[7,8] Combined with the solute and solvent reorganization energies, as well as the vibronic energy level splittings, these vibronic

[‡]Reproduced with permission from:
A. Sirjoosingh, and S. Hammes-Schiffer, "Diabatization Schemes for Generating Charge-Localized Electron-Proton Vibronic States in Proton-Coupled Electron Transfer Systems," *J. Chem. Theory Comput.* **7**, 2831-2841 (2011).

couplings can be used to calculate experimentally accessible quantities such as rate constants and kinetic isotope effects.

The objective of this paper is to develop the methodology for the rigorous construction of the charge-localized diabatic electron-proton vibronic states that form the basis of nonadiabatic PCET theories. A variety of schemes have been developed to generate diabatic electronic states, particularly for electron transfer reactions.[9-30] Diabatic electronic states can be defined mathematically as states with vanishing first-order nonadiabatic couplings at all possible nuclear configurations. In the context of electron transfer reactions, the diabatic electronic states are associated with the physically meaningful reactant and product states corresponding to the electron localized on the donor and acceptor, respectively. These diabatic electronic states are usually characterized by charge invariance in that the electronic charge distribution does not change significantly with nuclear motion. Schemes that have been developed to generate charge-localized diabatic electronic states for electron transfer reactions include the minimization of first-order nonadiabatic couplings using an adiabatic electronic state basis,[10-12] the generalized Mulliken-Hush method[13,14] and extensions using Boys localization,[24,25] block diagonalization methods,[19,20] approaches enforcing configurational uniformity,[9,16,17,22] constrained density functional theory,[28,29] and valence bond theory approaches.[15]

In a previous study,[23] we devised a scheme to calculate charge-localized diabatic electronic states for PCET reactions. In particular, we used an adiabatic-to-diabatic transformation[10] to generate diabatic electronic states constructed to ensure that the first-order nonadiabatic couplings with respect to the one-dimensional transferring hydrogen coordinate vanish exactly. When this diabatization approach was applied to the phenoxyl-phenol self-exchange PCET reaction, the diabatic electronic states were shown to exhibit physically

meaningful charge-localized electronic charge distributions. Moreover, we showed that this diabatization scheme provides quantitative diagnostics for the degree of electron-proton nonadiabaticity in PCET systems.[23] Identifying the degree of electron-proton nonadiabaticity is important because this property impacts the form of the vibronic coupling and the rate constant and provides insight into the fundamental mechanism. Specifically, electronically adiabatic proton transfer is associated with the hydrogen atom transfer mechanism, which does not involve significant electronic charge redistribution, and electronically nonadiabatic proton transfer is associated with the PCET mechanism, which involves significant electronic charge redistribution.[31,32]

In the present study, we expand this diabatization scheme in several directions. We extend this approach to general asymmetric PCET reactions and to other one-dimensional reaction coordinates, such as a normal mode coordinate or the intrinsic reaction coordinate (IRC). We also expand the previous treatment to describe three-dimensional hydrogen motion, where the component of the first-order nonadiabatic coupling vector along a specified one-dimensional reaction coordinate vanishes rigorously for all points on the three-dimensional diabatic electronic surfaces. In addition, we devise a strategy that utilizes the diabatic electronic states, along with the associated proton vibrational wavefunctions, to construct electron-proton vibronic states that describe either one-dimensional or three-dimensional hydrogen motion. These electron-proton vibronic states form the basis of nonadiabatic PCET theories and enable the calculation of vibronic couplings, rate constants, and kinetic isotope effects.

In addition to developing these extensions of the diabatization method for PCET reactions, we compare the diabatic electronic states obtained with this approach to those obtained with the generalized Mulliken-Hush (GMH) method[13,14] and extensions using Boys

localization.[24,25] These alternative diabatization methods generate the diabatic electronic states from the dipole moments associated with the ground and excited adiabatic electronic states rather than the first-order nonadiabatic couplings. The application of the GMH and Boys localization diabatization methods to PCET reactions is straightforward but, to our knowledge, has not been explored previously. The similarities among the diabatic electronic states generated with these three different diabatization methods provide a degree of validation for the underlying assumptions of the theoretical treatments.

An outline of this paper is as follows. Section 3.2A describes the adiabatic-to-diabatic transformation along a one-dimensional hydrogen coordinate utilized to generate diabatic electronic states. In Section 3.2B, we discuss modifications of this diabatization protocol to generate diabatic electronic states along general one-dimensional reaction coordinates, including a normal mode coordinate and the intrinsic reaction coordinate (IRC). Section 3.2C describes the construction of diabatic electron-proton vibronic states for both one-dimensional and three-dimensional hydrogen motion. Section 3.2D summarizes the GMH and Boys localization methods that are implemented for comparison. In Section 3.3, we provide the details of the computational methods used to study three model PCET systems: the phenoxyl-phenol self-exchange reaction, the asymmetric phenoxyl-quinol reaction, and the amidinium-carboxylate system representing an experimentally studied photoinduced PCET reaction.[33-35] Section 3.4A describes the generation of the diabatic electronic states for these three model systems, illustrating the extensions to asymmetric systems and to alternative one-dimensional reaction coordinates. Section 3.4B presents the strategy for combining the diabatic electronic states with the associated proton vibrational states to construct diabatic electron-proton vibronic states. In Section 3.4C, we provide a comparison of this diabatization method to the GMH and Boys

localization methods for generating diabatic electronic states. Finally, conclusions and future direction are discussed in Section 3.5.

3.2. Theory

A. Adiabatic-to-diabatic transformation

Consider a system comprised of N_e electrons, N_p protons, and N_s slow nuclei with coordinates $\mathbf{r}_e$, $\mathbf{r}_p$, and $\mathbf{R}$ and masses m_e, m_p and $\{M_I\}$, respectively, and with potential energy $V(\mathbf{r}_e, \mathbf{r}_p, \mathbf{R})$. The Hamiltonian for the "fast" degrees of freedom (i.e., the electron-proton subsystem) is

$$H_q = -\sum_{i'=1}^{N_p} \frac{\hbar^2}{2m_p} \nabla_{i'}^2 + H_e \,, \tag{3.1}$$

where the electronic Hamiltonian is

$$H_e = -\sum_{i=1}^{N_e} \frac{\hbar^2}{2m_e} \nabla_i^2 + V(\mathbf{r}_e, \mathbf{r}_p, \mathbf{R}). \tag{3.2}$$

For fixed $\mathbf{R}$, the eigenfunctions $\Phi_k(\mathbf{r}_e, \mathbf{r}_p; \mathbf{R})$ of H_q are calculated by solving

$$H_q \Phi_k(\mathbf{r}_e, \mathbf{r}_p; \mathbf{R}) = E_k(\mathbf{R}) \Phi_k(\mathbf{r}_e, \mathbf{r}_p; \mathbf{R}). \tag{3.3}$$

The adiabatic electronic states for fixed $(\mathbf{r}_p, \mathbf{R})$ are determined by solving

$$H_e \psi_i(\mathbf{r}_e; \mathbf{r}_p, \mathbf{R}) = \varepsilon_i(\mathbf{r}_p, \mathbf{R}) \psi_i(\mathbf{r}_e; \mathbf{r}_p, \mathbf{R}). \tag{3.4}$$

Assuming N electronic states, we define $\vec{\psi}$ to be a column vector of the N electronic eigenfunctions $\{\psi_i(\mathbf{r}_e; \mathbf{r}_p, \mathbf{R})\}$ of Eq. (3.4). Then we define an $N \times N$ transformation matrix $A(\mathbf{r}_p; \mathbf{R})$ such that

$$\vec{\xi}\left(\mathbf{r}_e;\mathbf{r}_p,\mathbf{R}\right)=A\left(\mathbf{r}_p;\mathbf{R}\right)\vec{\psi}\left(\mathbf{r}_e;\mathbf{r}_p,\mathbf{R}\right), \tag{3.5}$$

where $\vec{\xi}$ is a column vector of functions satisfying the condition

$$\left\langle \xi_i \left| \nabla_{\mathbf{r}_p} \xi_j \right\rangle_e = 0 \quad \text{for all } i, j. \right. \tag{3.6}$$

Thus, the transformed electronic states $\left\{\xi_i\left(\mathbf{r}_e;\mathbf{r}_p,\mathbf{R}\right)\right\}$ satisfy the standard definition of diabatic states with respect to the proton coordinate $\mathbf{r}_p$.

As discussed in Refs. 10 and 23, for a one-dimensional proton coordinate r_p and $N=2$ electronic states, the matrix A is given by

$$A\left(r_p;\mathbf{R}\right)=\begin{pmatrix}\cos\gamma & -\sin\gamma\\ \sin\gamma & \cos\gamma\end{pmatrix}, \tag{3.7}$$

where

$$\gamma\left(r_p;\mathbf{R}\right)=\gamma\left(r_0;\mathbf{R}\right)-\int_{r_0}^{r_p} d_{12}^{(ep)}\left(r;\mathbf{R}\right)\mathrm{d}r. \tag{3.8}$$

In this expression,

$$d_{12}^{(ep)}\left(r_p;\mathbf{R}\right)=\left\langle \psi_1 \left| \frac{\partial \psi_2}{\partial r_p} \right\rangle_e = \frac{\left\langle \psi_1 \left| \frac{\partial H_e}{\partial r_p} \right| \psi_2 \right\rangle_e}{\varepsilon_2-\varepsilon_1} \right. \tag{3.9}$$

is the first-order nonadiabatic coupling between adiabatic electronic states 1 and 2, and $\gamma\left(r_0;\mathbf{R}\right)$ is an additive constant that must be specified at some proton coordinate $r_p=r_0$. The diabatic potential energy matrix is given by $W=AUA^{-1}$, where $U_{ij}=\varepsilon_i\left(r_p,\mathbf{R}\right)\delta_{ij}$ is the adiabatic potential energy matrix. Here $W_{11}\left(r_p,\mathbf{R}\right)$ and $W_{22}\left(r_p,\mathbf{R}\right)$ are the diabatic electronic energies and $W_{12}\left(r_p,\mathbf{R}\right)$ is the diabatic electronic coupling.

Previously we applied this approach to symmetric systems. In this case, we chose $r_0 = 0$, corresponding to the transition state geometry, and set $\gamma(r_0) = -\pi/4$. This choice ensures that the adiabatic electronic states mix maximally and the diabatic electronic states cross at the transition state geometry, where the nonadiabatic coupling is a maximum. Moreover, the magnitude of the diabatic electronic coupling, W_{12}, is exactly half the splitting between the adiabatic electronic energies at this geometry. In the present paper, we extend this treatment to asymmetric systems, for which the nonadiabatic coupling is not necessarily a maximum at the transition state geometry. For the general case, we choose r_0 to be the hydrogen position at which the nonadiabatic coupling is maximum and set $\gamma(r_0) = -\pi/4$. This choice ensures that the diabatic states cross at $r_p = r_0$ and that the adiabatic states mix maximally at the hydrogen position corresponding to the largest nonadiabatic coupling.

B. Diabatization along other one-dimensional reaction coordinates

This approach may be extended to other one-dimensional reaction coordinates, such as a normal mode coordinate or an IRC. In this subsection, we discuss the generation of the diabatic electronic states along these types of alternative one-dimensional reaction coordinates, which typically are comprised of combinations of the motions of the transferring hydrogen and other heavy nuclei in the system. The objective is to calculate diabatic electronic states for which the first-order nonadiabatic coupling vanishes exactly along a general one-dimensional reaction coordinate. These diabatic electronic states will not be used to generate electron-proton vibronic states.

First we discuss the generation of diabatic electronic states along a single normal mode coordinate, q, with corresponding effective mass μ. In this case, we can still utilize Eqs. (3.1) to (3.4) by replacing r_p with q and m_p with μ, where the adiabatic and diabatic electronic energies depend explicitly on q rather than r_p. In PCET reactions, the relevant normal mode describing the proton transfer reaction is expected to be dominated by proton motion, so typically $\mu \approx m_p$. The adiabatic-to-diabatic transformation given in Eq. (3.5) ensures that the component of the nonadiabatic coupling vector along the normal mode coordinate q vanishes. Since the normal mode coordinate is a linear combination of Cartesian displacements of all nuclei, the nonadiabatic coupling with respect to the normal mode coordinate can be calculated analytically as a linear combination of the nonadiabatic couplings with respect to the Cartesian coordinates of all nuclei.

An alternative one-dimensional reaction coordinate is the IRC, which is generated numerically by following the minimum energy path from a transition state to the corresponding reactant and product state minima. In this case, the theoretical formalism described above is no longer rigorous, but we are able to define the adiabatic-to-diabatic transformation given in Eq. (3.5) so that the component of the first-order nonadiabatic coupling vector along the IRC vanishes. Since the IRC is generally not a linear combination of Cartesian coordinates, the nonadiabatic coupling in Eq. (3.9) cannot be calculated analytically. Instead, the component of the nonadiabatic coupling vector along the IRC can be calculated numerically as the scalar product of the nonadiabatic coupling vector with respect to the Cartesian coordinates of all nuclei and the instantaneous displacement vector of these coordinates with respect to the IRC approximated at each point using central-point differentiation. The formal treatment of the IRC in terms of Eqs. (3.1) - (3.4) is not rigorously valid because the IRC is not associated with a

specific mass and is not defined to have vanishing kinetic energy couplings with respect to other nuclear coordinates. As shown below, however, physically reasonable charge-localized diabatic electronic states for which the first-order nonadiabatic coupling vanishes along the IRC can be generated with this approach. Similarly, this approach may be used to generate these types of diabatic electronic states along any specified one-dimensional coordinate, such as the reaction path generated by a series of constrained optimizations, where the relative hydrogen position is constrained while all other nuclear coordinates are optimized.

C. Construction of electron-proton vibronic states

The diabatic electronic energies and couplings may be used to construct the electron-proton vibronic states that form the basis of nonadiabatic PCET theories. In the case of one-dimensional hydrogen motion with fixed heavy nuclei $\mathbf{R}$, the proton vibrational states for diabatic electronic state i are obtained by solving

$$\left(-\frac{\hbar^2}{2m_p}\frac{\partial^2}{\partial r_p^2}+W_{ii}\left(r_p,\mathbf{R}\right)\right)\varphi_\mu^{(i)}\left(r_p;\mathbf{R}\right)=\tilde{\varepsilon}_\mu^{(i)}\left(\mathbf{R}\right)\varphi_\mu^{(i)}\left(r_p;\mathbf{R}\right) \tag{3.10}$$

where $\tilde{\varepsilon}_\mu^{(i)}\left(\mathbf{R}\right)$ is the energy of the electron-proton vibronic state $\left(i,\mu\right)$. The diabatic electron-proton vibronic states $\left\{\zeta_{i\mu}\left(\mathbf{r}_e,r_p;\mathbf{R}\right)\right\}$ are then defined as products of the diabatic electronic wavefunctions and associated proton vibrational wavefunctions:

$$\zeta_{i\mu}\left(\mathbf{r}_e,r_p;\mathbf{R}\right)=\xi_i\left(\mathbf{r}_e;r_p,\mathbf{R}\right)\varphi_\mu^{(i)}\left(r_p;\mathbf{R}\right). \tag{3.11}$$

In the electronically nonadiabatic limit, the vibronic coupling between the reactant and product diabatic vibronic states $\zeta_{1\mu}$ and $\zeta_{2\nu}$ is[23,32,36]

$$V_{\mu\nu}^{(\mathrm{na})}=\left\langle\varphi_\mu^{(1)}\left|W_{12}\right|\varphi_\nu^{(2)}\right\rangle_p, \tag{3.12}$$

which reduces to the familiar form of the diabatic electronic coupling multiplied by the Franck-Condon overlap between the reactant and product proton vibrational wavefunctions when the diabatic electronic coupling $W_{12}\left(r_p,\mathbf{R}\right)$ is independent of r_p. In principle, a similar procedure could be applied to the diabatic electronic states generated along a normal mode coordinate when this normal mode is dominated by the hydrogen motion.

This treatment can be extended to construct three-dimensional electron-proton vibronic states that include the three-dimensional motion of the transferring hydrogen. In this case, the adiabatic and diabatic electronic states depend on the three-dimensional proton coordinate $\mathbf{r}_p=\left(r_p,s_p,t_p\right)$. The diabatic electronic states are constructed so that the component of the first-order nonadiabatic coupling vector along r_p, as given in Eq. (3.9), vanishes exactly for all points on a three-dimensional proton coordinate grid. Thus, the adiabatic-to-diabatic transformation can be expressed as

$$\vec{\xi}\left(\mathbf{r}_e;r_p,s_p,t_p,\mathbf{R}\right)=A\left(r_p,s_p,t_p;\mathbf{R}\right)\vec{\psi}\left(\mathbf{r}_e;r_p,s_p,t_p,\mathbf{R}\right), \tag{3.13}$$

but only the component of the first-order nonadiabatic coupling vector along r_p is used to determine $A\left(r_p,s_p,t_p;\mathbf{R}\right)$. In practice, the component of the first-order nonadiabatic coupling vector along r_p is calculated at each grid point $\left(r_p,s_p,t_p\right)$. For each value of $\left(s_p,t_p\right)$, r_0 is chosen to be the proton position along the r_p slice at which this nonadiabatic coupling is maximum, and $\gamma\left(r_p,s_p,t_p\right)$ is determined by calculating the line integral in Eq. (3.8) with $\gamma\left(r_0,s_p,t_p\right)=-\pi/4$. This procedure ensures that the first-order nonadiabatic couplings with respect to r_p vanish exactly for all points on the three-dimensional proton coordinate grid.

The resulting three-dimensional diabatic potential energy matrix is given by $W = AUA^{-1}$, where $U_{ii} = \varepsilon_i\left(r_p, s_p, t_p, \mathbf{R}\right)$ are the three-dimensional adiabatic potential energy surfaces and $W_{ii}\left(r_p, s_p, t_p, \mathbf{R}\right)$ are the three-dimensional diabatic potential energy surfaces. For fixed heavy nuclei **R**, the three-dimensional proton vibrational wavefunctions for diabatic electronic state i are calculated using the three dimensional analog of Eq. (3.10) with potential energies $W_{ii}\left(r_p, s_p, t_p, \mathbf{R}\right)$. These wavefunctions can then be combined with the associated diabatic electronic states, as in Eq. (3.11), to form three-dimensional electron-proton vibronic states.

We also explored an alternative more approximate formulation in which the adiabatic-to-diabatic transformation matrix A defined in Eq. (3.7) depends explicitly on only the one-dimensional proton coordinate r_p [i.e., $\gamma\left(r_p\right)$ is independent of the orthogonal coordinates (s_p, t_p)]. The advantage of this alternative formulation is that it only requires the calculation of the first-order nonadiabatic couplings along the one-dimensional r_p axis corresponding to $s_p = t_p = 0$ (i.e., the proton donor-acceptor axis in the applications discussed below). As a result, however, the first-order nonadiabatic couplings with respect to r_p vanish only along the one-dimensional r_p axis and do not vanish exactly for the other points on the three-dimensional proton coordinate grid.

D. Generalized Mulliken-Hush and Boys localization

The GMH method[13,14] generates diabatic electronic states using the adiabatic electronic state dipole moments. These dipole moments are defined in terms of the adiabatic electronic

states as $\vec{\mu}_{ij} = \langle \psi_i | \hat{\mu} | \psi_j \rangle$ for $i \in \{1,2\}$, where $\hat{\mu}$ is the dipole moment operator. The GMH method utilizes an adiabatic-to-diabatic transformation matrix analogous to Eq. (3.7):

$$A^{\mathrm{GMH}}\left(r_p;\mathbf{R}\right) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{3.14}$$

In this case, $\theta\left(r_p;\mathbf{R}\right)$ depends on the adiabatic electronic state dipole moments:

$$\tan\left(2\theta\right) = \frac{2\vec{\mu}_{12}\cdot\vec{v}}{\left|\vec{v}\right|^2}, \tag{3.15}$$

where $\vec{v} = \vec{\mu}_{11} - \vec{\mu}_{22}$ and typically $\vec{\mu}_{12}$ and $\vec{v}$ are assumed to be parallel. Analogous to Eq. (3.5), the diabatic electronic states are determined by

$$\vec{\xi}^{\mathrm{GMH}}\left(\mathbf{r}_e;r_p,\mathbf{R}\right) = A^{\mathrm{GMH}}\left(r_p;\mathbf{R}\right)\vec{\psi}\left(\mathbf{r}_e;r_p,\mathbf{R}\right). \tag{3.16}$$

The GMH expression for the mixing angle in Eq. (3.15) is derived by defining A^{GMH} as the transformation that diagonalizes the adiabatic dipole moment matrix, ensuring that the diabatic transition dipole moment is exactly zero, i.e., $\left\langle \xi_1^{\mathrm{GMH}} \mid \hat{\mu} \mid \xi_2^{\mathrm{GMH}} \right\rangle = 0$.

Subotnik *et al*. demonstrated that the GMH procedure can be extended to arbitrary geometries and multiple charge centers using Boys localization.[25] The more general expression for θ in Eq. (3.14) is given as[25]

$$\cos\left(4\theta\right) = \frac{-F}{\sqrt{F^2+G^2}}, \tag{3.17}$$

where

$$F = \left|\vec{\mu}_{12}\right|^2 - \frac{\left|\vec{\mu}_{11}-\vec{\mu}_{22}\right|^2}{4} \tag{3.18}$$

and

$$G = \vec{\mu}_{12}\cdot\left(\vec{\mu}_{11}-\vec{\mu}_{22}\right). \tag{3.19}$$

Invoking the assumption that $\vec{\mu}_{12}$ and $\vec{v}$ are parallel in Eqs. (3.17) - (3.19) leads to an expression for θ that coincides with the GMH expression given in Eq. (3.15).

3.3. Computational Methods

We used three systems to test the theoretical methods developed above. We studied the self-exchange reaction between the phenoxyl radical and the phenol molecule depicted in Figure 3.1a. We also studied the analogous reaction between the phenoxyl radical and the 1,4-benzenediol (quinol) molecule depicted in Figure 3.1b. The transition state geometries for these systems were calculated using density functional theory (DFT) with the B3LYP functional[37,38] and the 6-31G* basis set.[39] We also examined the amidinium-carboxylate system depicted in Figure 3.1c. We followed a similar but not identical geometry optimization procedure as that described in Ref. 35. In our procedure, we optimized the complete neutral amidinium-carboxylate system at the RHF/6-31G** level of theory, maintaining planarity of the system except for the three methyl group hydrogen atoms. The adiabatic and diabatic electronic states were calculated for the PCET reaction in the negatively charged complex with the lower proton in Figure 3.1c being transferred. The geometry optimizations were performed using Gaussian 09.[40]

We generated the adiabatic and diabatic electronic energy curves for the phenoxyl-phenol and phenoxyl-quinol systems along three different types of one-dimensional reaction coordinates. In the first scheme,[23] all nuclei except the transferring hydrogen atom were fixed at the transition state geometry, and the hydrogen was displaced along a one-dimensional grid spanning the hydrogen donor-acceptor axis. In the second scheme, all nuclei were displaced according to the normal mode coordinate corresponding to the negative frequency identified at

the transition state geometry. In the third scheme, all nuclei were displaced according to the IRC calculated using the same level of theory as described above. For all three schemes, we generated 128 geometries along the reaction coordinate. We used only the first scheme, which is based on the one-dimensional hydrogen coordinate along the hydrogen donor-acceptor axis with all other nuclei fixed, to study the amidinium-carboxylate system.

For each system, we calculated the two lowest-energy electronically adiabatic potential energy curves using the complete active space self-consistent field (CASSCF) method. The CASSCF calculations of the phenoxyl-phenol and phenoxyl-quinol systems were performed with the 6-31G* basis set and an active space of three electrons in six orbitals, state-averaging over the ground and first excited electronic states with equal weighting. The CASSCF calculations of the amidinium-carboxylate system were performed with the 6-31G** basis set and an active space of one electron in three orbitals, state-averaging over the lowest three electronic states with equal weighting.[35] We also calculated the two lowest-energy electronically adiabatic potential energy surfaces in three dimensions by displacing the hydrogen on a three-dimensional grid consisting of 16 points in each direction, with all other nuclei remaining fixed. In addition, we obtained the nonadiabatic coupling vectors with respect to the Cartesian coordinates of all nuclei directly from the CASSCF calculations. The GAMESS electronic structure package[41] was used to perform all CASSCF calculations. Note that these calculations were performed at a relatively low level of theory because our goal is to examine only the qualitative features of the various theoretical approaches.

We constructed the electron-proton vibronic states for one-dimensional or three-dimensional hydrogen motion with all other nuclei fixed. This procedure utilized the diabatic electronic states generated with the first scheme described above. We calculated one-dimensional

proton vibrational wavefunctions describing the proton motion on the diabatic electronic energy surfaces by solving the one-dimensional Schrödinger equation given in Eq. (3.10). These calculations were performed with the Fourier grid Hamiltonian (FGH) method[42] using 128 grid points along the hydrogen donor-acceptor axis. We calculated three-dimensional proton vibrational wavefunctions using the three-dimensional analog of Eq. (3.10). These calculations were performed with the FGH-FCI (full configuration interaction) method[43] on the three-dimensional diabatic electronic energy surfaces mentioned above.

To examine the charge transfer properties, we calculated the dipole moments, atomic charges, and electrostatic potential maps as functions of the one-dimensional reaction coordinate for the adiabatic and diabatic electronic states. The properties of the adiabatic electronic states were calculated directly from the CASSCF wavefunctions with GAMESS. For the diabatic electronic states, we modified a local version of GAMESS to calculate these properties for the appropriate linear combination of configuration interaction states following the transformation given in Eq. (3.5). The atomic charges were obtained by fitting to the electrostatic potential calculated at points on the Connolly surface[44] under the constraint of reproducing the total charge and dipole moment of the electronic state under consideration.[45]

3.4. Results

A. Generating diabatic electronic states

In this study, we applied the adiabatic-to-diabatic transformation methodology[23] to the more general case of asymmetric PCET reactions, as depicted in Figures 3.1b and 3.1c. For each system, we calculated the ground and first excited adiabatic electronic states and the nonadiabatic couplings along the one-dimensional hydrogen coordinate using the CASSCF

method. As shown in Figures 3.2 and 3.3, the adiabatic-to-diabatic transformation successfully generated physically meaningful charge-localized diabatic electronic states for all three systems. In all cases, the diabatic electronic energies, $W_{11}\left(r_p,\mathbf{R}\right)$ and $W_{22}\left(r_p,\mathbf{R}\right)$ (dashed blue and red lines, respectively, in the left panels of Figure 3.2) are virtually identical to the adiabatic electronic energies, $\varepsilon_1\left(r_p,\mathbf{R}\right)$ and $\varepsilon_2\left(r_p,\mathbf{R}\right)$ (solid black lines in the left panels of Figure 3.2) over all hydrogen positions except near $r_p = r_0$, where they smoothly cross. All three systems exhibit a relatively localized region of strong nonadiabatic coupling, as shown in the right panels of Figure 3.2.

Figures 3.2b and 3.2c illustrate that the phenoxyl-quinol and the amidinium-carboxylate systems possess significantly asymmetric adiabatic electronic energies and nonadiabatic couplings. In particular, these systems exhibit an energy bias between geometries corresponding to the hydrogen localized on the donor molecule and the hydrogen localized on the acceptor molecule. Furthermore, the maximum of the nonadiabatic coupling along the donor-acceptor axis is shifted from the donor-acceptor midpoint, which is chosen to be at $r_p = 0$ for all systems. This shift is more noticeable for the phenoxyl-quinol system. As mentioned above, we chose r_0 to be the hydrogen position at which the nonadiabatic coupling is maximum and set $\gamma\left(r_0\right) = -\pi/4$ in Eq. (3.8) to ensure that the diabatic states cross at $r_\mathrm{p} = r_0$ and that the adiabatic states mix maximally at the hydrogen position corresponding to the largest nonadiabatic coupling. Figure 3.2 illustrates that the resulting diabatic potential energies correspond to the transferring hydrogen localized on the donor molecule in $W_{11}\left(r_p,\mathbf{R}\right)$ and the acceptor molecule in $W_{22}\left(r_p,\mathbf{R}\right)$.

The electrostatic potential maps for diabatic states ξ_1 and ξ_2 at $r_p = r_0$ are depicted in Figure 3.3. These electrostatic potential maps illustrate that the diabatic electronic states possess localized electronic charge distributions. The corresponding electrostatic potential maps for other positions of the transferring hydrogen are provided in Appendix B (Figures B.1, B.2, and B.3) and indicate that these electronic charge distributions are relatively invariant along the transferring hydrogen coordinate. The amidinium-carboxylate system is negatively charged, and the electrostatic potential maps are strongly influenced by the charge separation at the hydrogen-bonding interface (see Figure 3.1c). To clarify the charge localization of the diabatic electronic states with respect to the transferring electron, the electrostatic potential of the neutral amidinium-carboxylate complex was subtracted from that of the negatively charged complex. The resulting difference electrostatic potential maps plotted in Figure 3.3 clearly demonstrate charge localization with respect to the transferring electron.

Figure 3.4a and Figure 3.5 depict the partial charges on the donor (green) and acceptor (purple) molecules for the adiabatic (left panels) and diabatic (center and right panels) electronic states for the phenoxyl-phenol and amidinium-carboxylate systems, respectively. The corresponding figure for the phenoxyl-quinol system is provided in Figure B.4. For all three systems, the partial charges on the donor and acceptor molecules change significantly along the reaction coordinate for the adiabatic electronic states but remain relatively constant for the diabatic electronic states. Analogous to the procedure used for the electrostatic potential maps, the partial charges for the donor and acceptor molecules in the amidinium-carboxylate system are determined by subtracting the corresponding partial charges calculated for the neutral complex. The partial charges prior to this subtraction are provided in Figure B.5. As shown in Figure 3.5, the resulting plots clearly demonstrate the charge localization of the diabatic

electronic states with respect to the transferring electron. Thus, this adiabatic-to-diabatic transformation method provides physically meaningful diabatic electronic states with localized electronic charge distributions that are relatively invariant along the transferring hydrogen coordinate for both symmetric and asymmetric systems.

We also calculated the diabatic electronic states along two other types of reaction coordinates: the normal mode coordinate associated with the negative frequency at the transition state and the IRC. We present the results for the phenoxyl-phenol system, although the corresponding results for the asymmetric phenoxyl-quinol system are qualitatively similar to those for the symmetric case and are included in Figure B.6. Figure 3.6 depicts the adiabatic and diabatic electronic energies calculated along the normal mode coordinate and the IRC. We observed that the normal mode coordinate is dominated by the motion of the transferring hydrogen (i.e., the mass associated with this normal mode was ≈1.12 amu). For this reason, the adiabatic and diabatic electronic energies calculated along the normal mode coordinate (Figure 3.6a) are very similar to those calculated along the one-dimensional hydrogen coordinate (Figure 3.2a). The adiabatic and diabatic electronic energies calculated along the IRC (Figure 3.6b) are qualitatively similar to those calculated along the one-dimensional hydrogen coordinate near the transition state but plateau in the outer regions after the IRC reaches the minimum energy geometries.

Figure 3.4 compares the partial charges on the donor (green) and acceptor (purple) molecules for the adiabatic (left panels) and diabatic (center and right panels) electronic states for the three different types of reaction coordinates. The partial charges for the one-dimensional hydrogen coordinate and the normal mode coordinate, as depicted in Figures 3.4a and 3.4b, respectively, are very similar. The results for the IRC are also qualitatively similar. In all cases,

the partial charges on the donor and acceptor molecules remain relatively constant along the reaction coordinate for the diabatic electronic states. Thus, the adiabatic-to-diabatic transformation method provides charge-localized diabatic electronic states with electronic charge distributions that are relatively invariant along all three types of reaction coordinates.

B. Constructing electron-proton vibronic states

We constructed the electron-proton vibronic states corresponding to the diabatic electronic states generated along the one-dimensional hydrogen coordinate with all other nuclei fixed. For this purpose, we calculated the one-dimensional proton vibrational wavefunctions by solving Eq. (3.10) with each of the diabatic electronic energies, $W_{11}\left(r_p,\mathbf{R}\right)$ and $W_{22}\left(r_p,\mathbf{R}\right)$, shown in Figure 3.2a. The four lowest-energy proton vibrational wavefunctions corresponding to each diabatic electronic state for the phenoxyl-phenol system are depicted in Figure 3.7. These proton vibrational wavefunctions can be combined with the diabatic electronic wavefunctions, ξ_1 and ξ_2, by forming products as in Eq. (3.11) to obtain the electron-proton vibronic states that comprise the basis of nonadiabatic PCET rate theories. In particular, these vibronic states can be used directly to calculate vibronic couplings, which are essential for the calculation of experimentally accessible quantities such as rate constants and kinetic isotope effects. The vibronic couplings calculated using Eq. (3.12) will be discussed below.

We also calculated three-dimensional diabatic potential energy surfaces and the associated proton vibrational wavefunctions. For this purpose, we applied the adiabatic-to-diabatic transformation given by Eq. (3.13), where the transformation matrix depends on the three-dimensional proton coordinate, and the component of the nonadiabatic coupling vector along r_p vanishes exactly for all points on the three-dimensional proton coordinate grid. The

resulting partial charges for the phenoxyl-phenol system are given in Table 3.1. For comparison, we also applied the more approximate adiabatic-to-diabatic transformation, where the transformation matrix depends explicitly on only the one-dimensional proton coordinate r_p, and the component of the nonadiabatic coupling vector along r_p vanishes exactly only for points on the one-dimensional proton donor-acceptor axis. The resulting partial charges are given in Table B.1 and are qualitatively similar to those provided in Table 3.1. This more approximate approach may be useful for larger systems because the computational expense is significantly lower.

Table 3.1 demonstrates that the three-dimensional diabatic electronic states generated for the phenoxyl-phenol system are charge-localized. This table provides the partial charges on the donor and acceptor molecules, as well as the transferring hydrogen, averaged over all hydrogen positions on the three-dimensional grid. As expected, the average donor and acceptor partial charges for the three-dimensional adiabatic electronic ground state are identical, and the standard deviations reflect the changes in the donor and acceptor partial charges that are consistent with those observed along the one-dimensional hydrogen coordinate depicted in the left panel of Figure 3.4a. In contrast, the three-dimensional diabatic electronic states exhibit charge-localization: the average charge on the donor is negative (positive) and the average charge on the acceptor is positive (negative) when averaged over all hydrogen positions for the diabatic electronic state ξ_1 (ξ_2). These results are also consistent with the partial charges along the one-dimensional hydrogen coordinate depicted in the center and right panels of Figure 3.4a. The standard deviations are slightly larger than would be predicted from Figure 3.4 because of numerical fluctuations near the transition state geometry.

We used these three-dimensional diabatic potential energy surfaces to calculate the associated three-dimensional proton vibrational wavefunctions with the FGH-FCI method. The three lowest-energy proton vibrational wavefunctions corresponding to each diabatic electronic state for the phenoxyl-phenol system are depicted in Figure 3.8. The proton vibrational wavefunctions associated with the diabatic state ξ_1 (blue) are localized near the donor molecule, while the proton vibrational wavefunctions associated with the diabatic state ξ_2 (red) are localized near the acceptor molecule. These results are consistent with the corresponding one-dimensional proton vibrational wavefunctions along the donor-acceptor axis depicted in Figure 3.7. The three-dimensional proton vibrational wavefunctions can be combined with the corresponding diabatic electronic wavefunctions to generate three-dimensional diabatic electron-proton vibronic states, which can be used to calculate vibronic couplings, rate constants and kinetic isotope effects of nonadiabatic PCET reactions.

C. Comparison to generalized Mulliken-Hush and Boys localization

We also used the GMH procedure to generate charge-localized diabatic electronic states for the phenoxyl-phenol and amidinium-carboxylate systems. The adiabatic state dipole moments were calculated directly from the CASSCF wavefunctions. The GMH diabatization produced qualitatively similar diabatic electronic states to those generated using the diabatization procedure described in Section 3.2A. In particular, the diabatic electronic energies depicted in Figures 3.9a and 3.9b are virtually identical to those depicted in Figures 3.2a and 3.2c, respectively. Moreover, Figure 3.10 illustrates that the calculated diabatic state dipole moments are very similar for the two diabatization schemes. Thus, these two fundamentally different diabatization approaches lead to nearly identical diabatic electronic states. In principle, proton

vibrational wavefunctions could be calculated using the GMH diabatic potential energies and combined with the associated diabatic electronic wavefunctions to form electron-proton vibronic states, as described above.

In addition, we used the Boys localization scheme to generate diabatic electronic states for the phenoxyl-phenol and amidinium-carboxylate systems. The Boys localization method produced results that are virtually identical to the GMH results presented in Figures 3.9 and 3.10. This high level of agreement is attributed to the observation that the vectors $\vec{\mu}_{12}$ and $\vec{\mu}_{11} - \vec{\mu}_{22}$ were essentially parallel for all positions of the transferring hydrogen. In particular, for all r_p such that $|r_p| \leq 0.5$ Angstroms, $\left|(\vec{\mu}_{12})^u \cdot (\vec{\mu}_{11} - \vec{\mu}_{22})^u\right| \geq 0.998$ for the phenoxyl-phenol system and $\left|(\vec{\mu}_{12})^u \cdot (\vec{\mu}_{11} - \vec{\mu}_{22})^u\right| \geq 0.992$ for the amidinium-carboxylate system, where the *u* superscript indicates the unit vector in the specified direction. Thus, the GMH and Boys localization methods are in excellent agreement with the diabatization method presented in Section 3.2A for the systems studied.

Finally, we calculated the electronic coupling at the geometry corresponding to the crossing point of the diabatic potential energy curves for the three systems studied. This quantity, $W_{12}(r_0; \mathbf{R})$, is calculated as the off-diagonal element of the diabatic potential energy matrix obtained from transforming the adiabatic potential energy matrix at $r_p = r_0$. Table 3.2 demonstrates that the electronic couplings calculated using the GMH and Boys localization methods agree very well with those obtained from the diabatization procedure described in Section 3.2A. Table 3.2 also provides the vibronic couplings between the ground electron-proton vibronic states calculated using Eq. (3.12) for μ = ν = 0. These vibronic couplings are

significantly smaller than the corresponding electronic couplings because of the relatively small overlap between the reactant and product ground state proton vibrational wavefunctions.

3.5. Conclusions

In this paper, we developed a scheme to generate charge-localized diabatic electronic states for a wide range of PCET systems. These charge-localized diabatic electronic states are obtained from standard electronic structure calculations using an adiabatic-to-diabatic transformation designed to ensure that the first-order nonadiabatic couplings with respect to a specified one-dimensional reaction coordinate vanish exactly. We applied this protocol to both symmetric and asymmetric PCET systems with several different one-dimensional reaction coordinates, including the hydrogen transfer coordinate, a normal mode coordinate, and the IRC. This approach was also extended to construct three-dimensional charge-localized diabatic electronic surfaces corresponding to the three-dimensional motion of the transferring hydrogen. We demonstrated that this methodology leads to physically meaningful charge-localized diabatic electronic states with relatively invariant charge distributions along the reaction coordinate. These diabatic electronic states are in excellent agreement with those obtained from the GMH and Boys localization methods.

In addition, we combined these diabatic electronic states with the associated proton vibrational wavefunctions to generate electron-proton vibronic states that describe one- or three-dimensional hydrogen motion. These electron-proton vibronic states can be used to calculate the vibronic couplings that enter the nonadiabatic rate constant expressions for PCET reactions. Within the golden rule formalism, each term in the nonadiabatic PCET rate constant expression is proportional to the square of the electron-proton vibronic coupling for a pair of reactant and

product vibronic states. As a result, the vibronic couplings strongly impact the rate constants and kinetic isotope effects of PCET reactions. Thus, the construction of charge-localized electron-proton vibronic states is essential for the calculation of experimentally measurable quantities such as the rate constants and kinetic isotope effects of PCET reactions.

Electronic State	Donor Charge	Acceptor Charge	H Charge
ψ_1	–0.2±0.4	–0.2±0.4	0.5±0.1
ξ_1	–0.7±0.2	0.3±0.2	0.4±0.1
ξ_2	0.3±0.2	–0.7±0.2	0.4±0.1

Table 3.1: Average electrostatic potential-derived partial charges calculated for the three-dimensional ground adiabatic and diabatic electronic states reported with standard deviations for the phenoxyl-phenol system. The tabulated results were computed by averaging over the 16^3 hydrogen positions on the three-dimensional grid, where points around $r_p = r_0$ were omitted due to numerical noise in this region. All charges are given in units of e. Deviation of the total charge from zero is due to numerical noise and round-off error. The diabatic electronic states were calculated using Eq. (3.13), ensuring that the component of the first-order nonadiabatic coupling vector along r_p vanishes exactly for all points on the three-dimensional grid.

System	V_{el}^{GMH}	V_{el}^{Boys}	V_{el}	$V_{00}^{(na)}$
phenol	606	606	606	7.2
quinol	610	610	611	8.7
amidinium-carboxylate	205	205	206	0.44

Table 3.2: Electronic and vibronic couplings calculated at the geometry corresponding to the crossing point of the diabatic potential energy curves using the various diabatization methods. The electronic couplings were calculated as the off-diagonal element of the diabatic potential energy matrix using the GMH method, the Boys localization method, and the method described in Section 3.2A for V_{el}^{GMH}, V_{el}^{Boys}, and V_{el}, respectively. The vibronic couplings $V_{00}^{(na)}$ between ground electron-proton vibronic states were calculated with V_{el} using Eq. (3.12) for $\mu = \nu = 0$. All couplings are given in units of cm^{-1}.

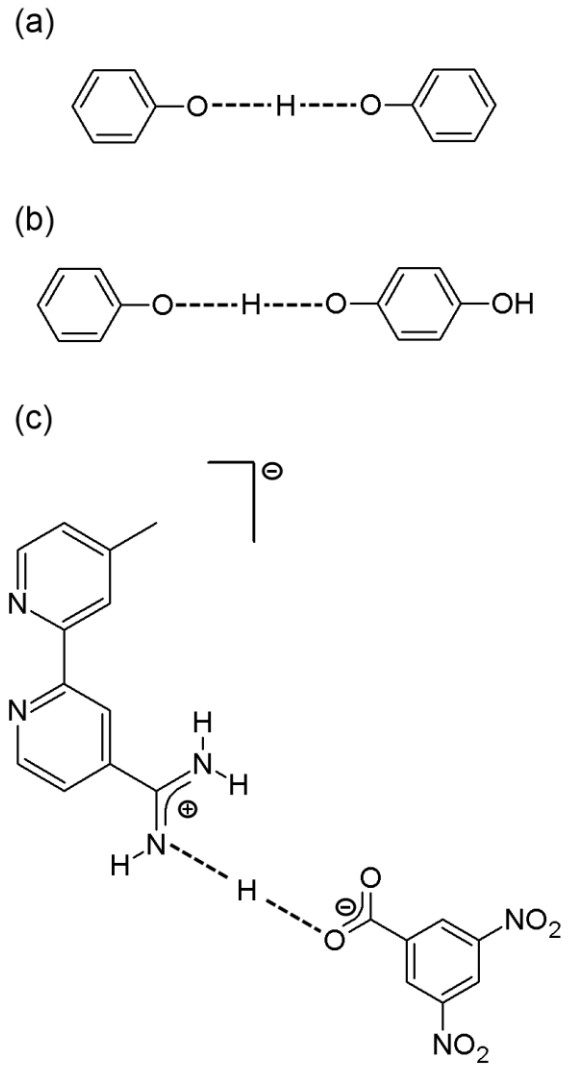

Figure 3.1: Three model systems studied: (a) phenoxyl-phenol, (b) phenoxyl-quinol and (c) amidinium-carboxylate systems. The phenoxyl-phenol and phenoxyl-quinol systems are neutral, while the amidinium-carboxylate system has an overall charge of −1.

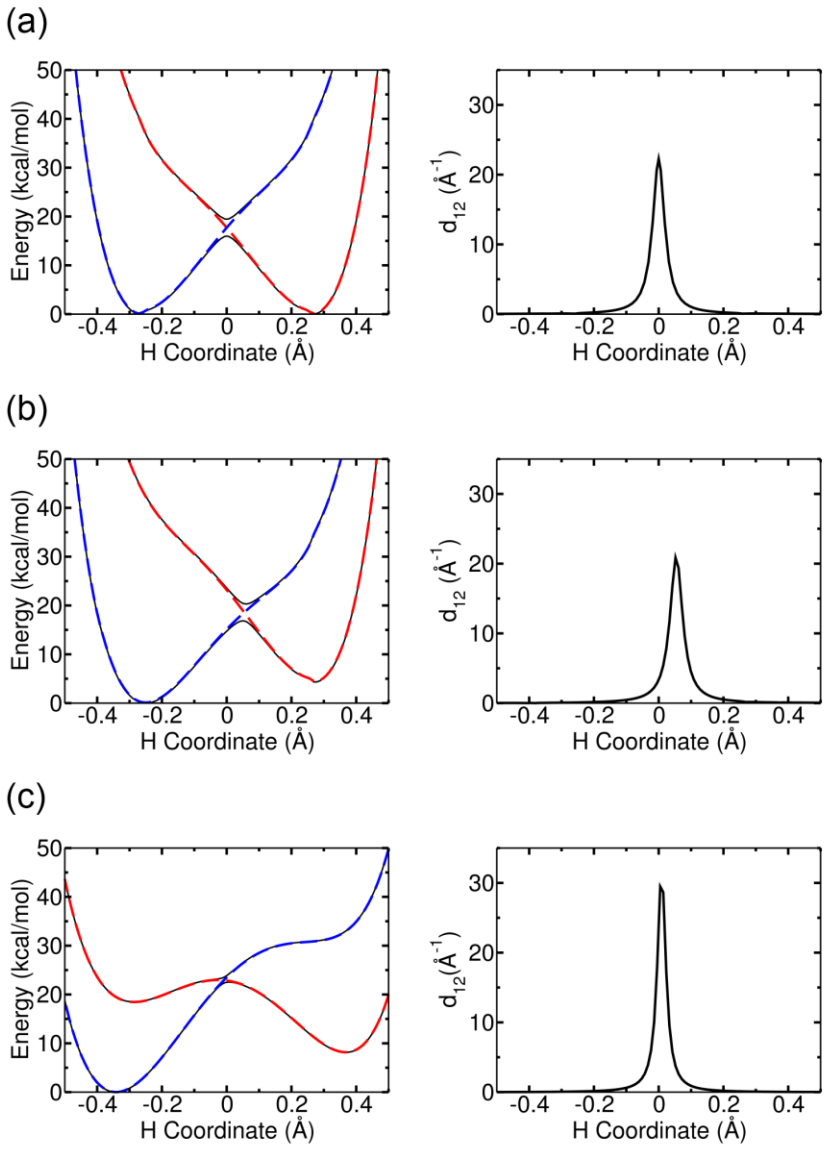

Figure 3.2: Adiabatic and diabatic electronic state properties as functions of the transferring hydrogen coordinate for the (a) phenoxyl-phenol, (b) phenoxyl-quinol and (c) amidinium-carboxylate systems. The left panels depict the electronically adiabatic and diabatic potential energy curves. The solid black curves are the ground and first excited adiabatic state energies $\varepsilon_1\left(r_p,\mathbf{R}\right)$ and $\varepsilon_2\left(r_p,\mathbf{R}\right)$, respectively, calculated with the CASSCF method, and the dashed blue and red curves are the diabatic electronic energies $W_{11}\left(r_p,\mathbf{R}\right)$ and $W_{22}\left(r_p,\mathbf{R}\right)$, respectively, where the choice of $\gamma\left(r_0\right)$ is described in the text. The right panels depict the component of the first-order nonadiabatic coupling vector along the hydrogen donor-acceptor axis, as defined in Eq. (3.9).

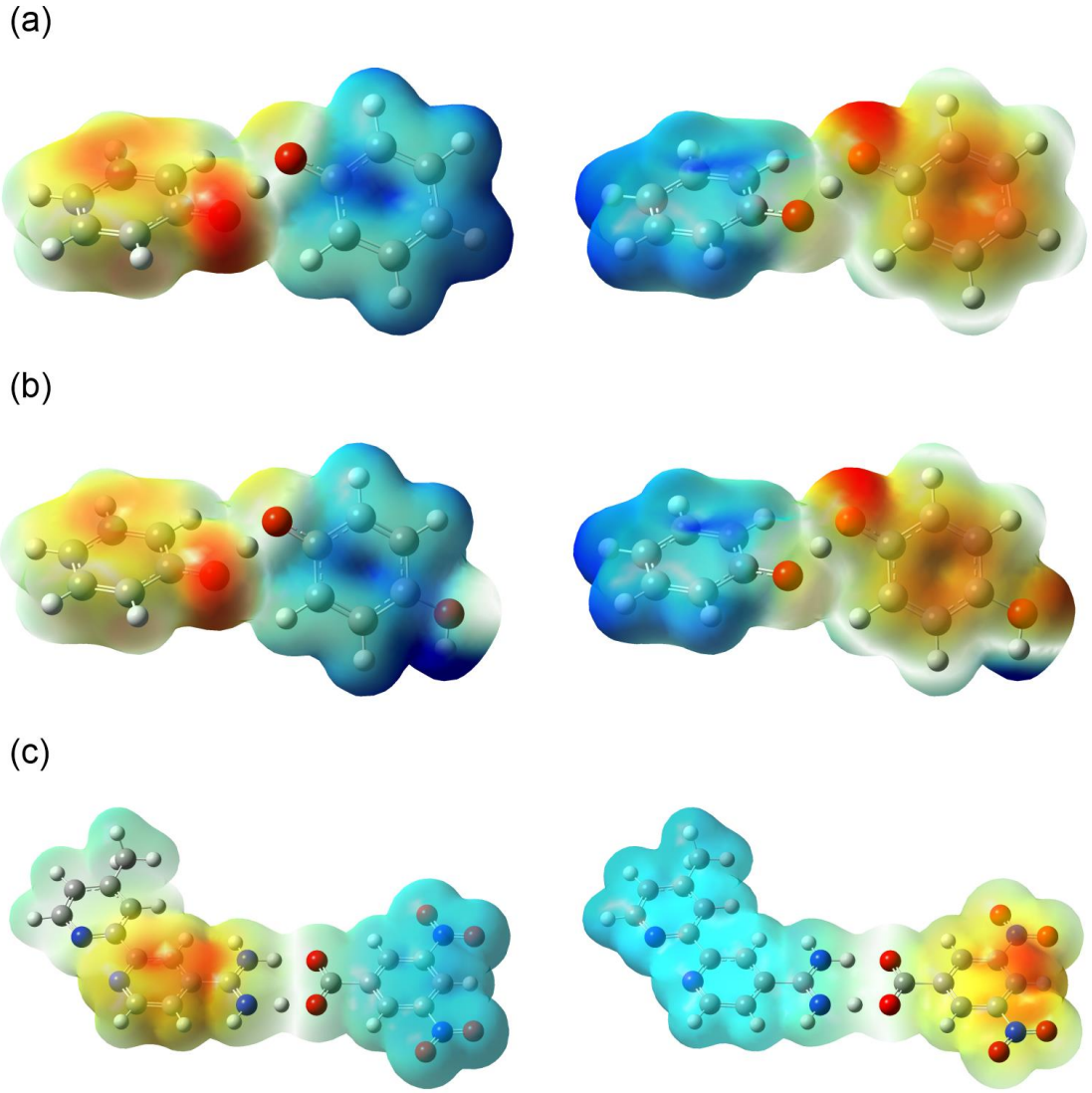

Figure 3.3: Electrostatic potential maps for the diabatic electronic states ξ_1 (left) and ξ_2 (right) corresponding to a density isosurface value of 0.005 for the (a) phenoxyl-phenol, (b) phenoxyl-quinol and (c) amidinium-carboxylate systems at $r_p = r_0$. Negatively and positively charged regions are indicated by red and blue coloring, respectively. The maps for the amidinium-carboxylate system are difference electrostatic potential maps with respect to the neutral complex, as described in the text.

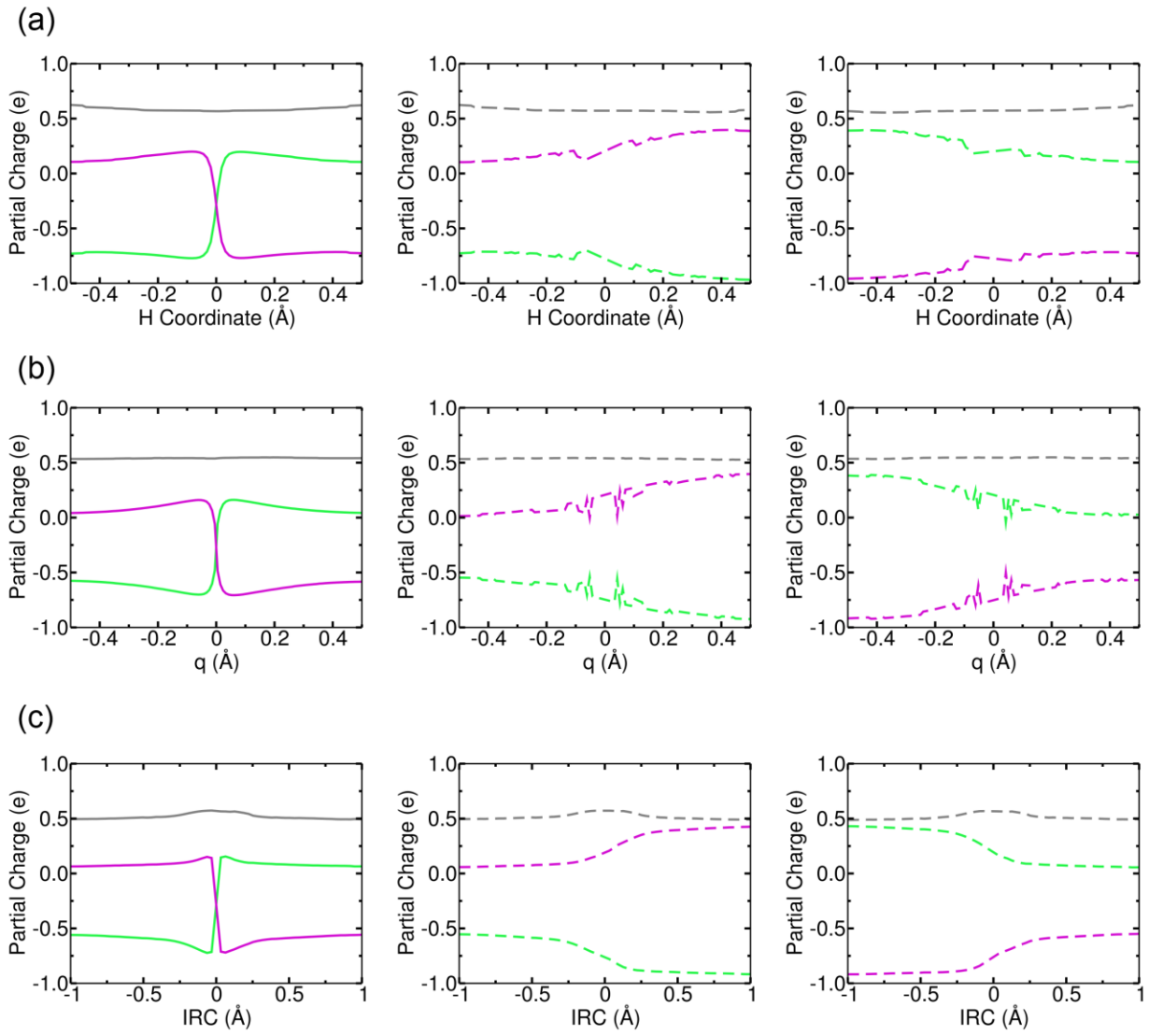

Figure 3.4: Partial charges determined from electrostatic potential-derived atomic charges for the ground adiabatic electronic state (left), the diabatic electronic state ξ_1 (center) and the diabatic electronic state ξ_2 (right) for the phenoxyl-phenol system calculated using various reaction coordinates: (a) the one-dimensional hydrogen coordinate, (b) the normal mode coordinate corresponding to the negative frequency at the transition state geometry, and (c) the IRC. Partial charges are shown for the donor molecule (green), acceptor molecule (purple), and transferring hydrogen (gray). Calculated values of the partial charges for the diabatic electronic states around $r_p = r_0$ are omitted due to numerical noise in this region.

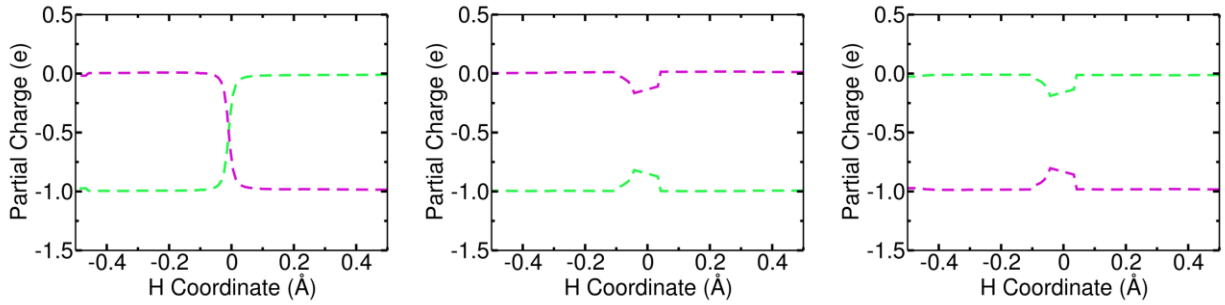

Figure 3.5: Partial charges determined from electrostatic potential-derived atomic charges for the ground adiabatic electronic state (left), the diabatic electronic state ξ_1 (center) and the diabatic electronic state ξ_2 (right) as functions of the transferring hydrogen coordinate for the amidinium-carboxylate system. Partial charges are shown for the donor molecule (green) and the acceptor molecule (purple). The partial charges on the donor and acceptor molecules are obtained after subtracting the corresponding partial charges of the neutral complex, as described in the text. Due to this subtraction, the transferring hydrogen has no significant charge, although it has a nearly constant charge of ~ 0.58 e for the adiabatic and diabatic electronic states prior to this subtraction. Calculated values of the partial charges for the diabatic electronic states around $r_p = r_0$ are omitted due to numerical noise in this region.

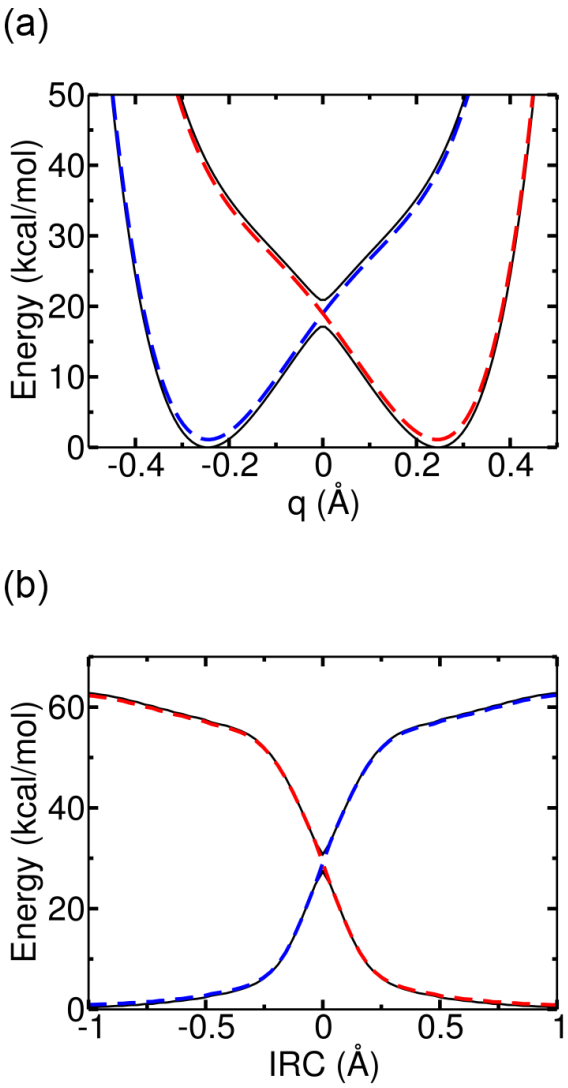

Figure 3.6: Electronically adiabatic and diabatic potential energy curves for the phenoxyl-phenol system as functions of (a) the normal mode coordinate corresponding to the negative frequency at the transition state geometry and (b) the IRC. The solid black curves are the ground and first excited adiabatic state energies, and the dashed blue and red curves are the diabatic electronic energies.

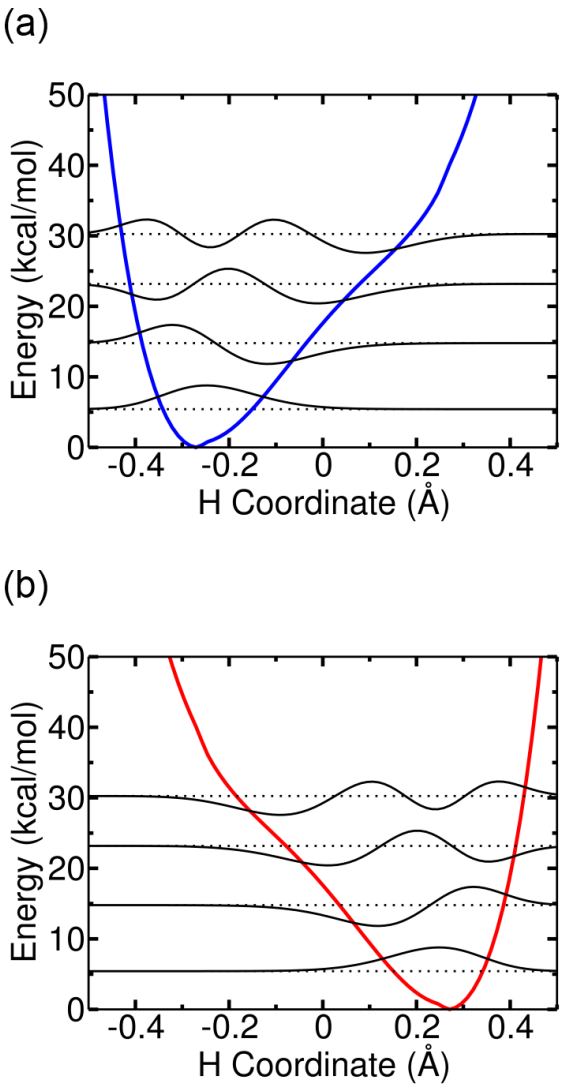

Figure 3.7: Four lowest-energy one-dimensional proton vibrational wavefunctions (black solid lines) calculated using Eq. (3.10) for the diabatic electronic potential (a) W_{11} and (b) W_{22} for the phenoxyl-phenol system.

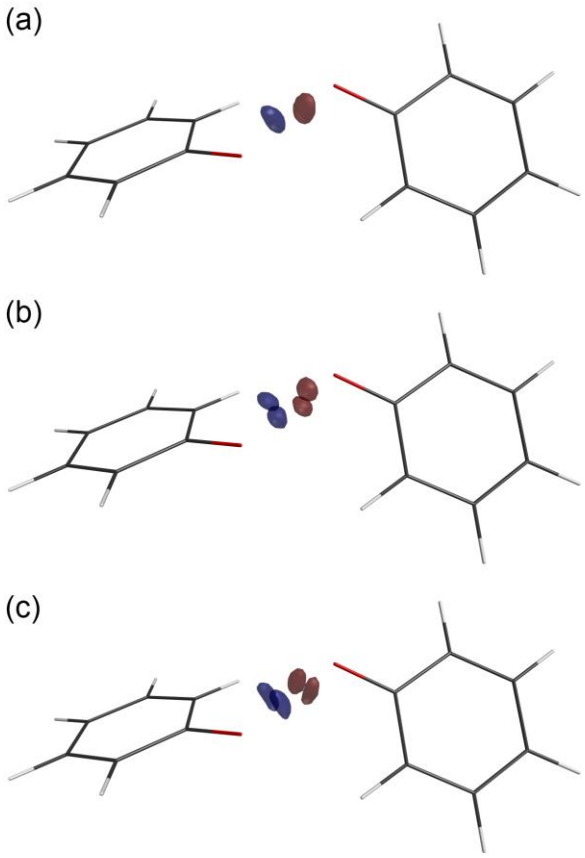

Figure 3.8: Three lowest-energy three-dimensional proton vibrational wavefunctions corresponding to a density isosurface value of 0.002 for the three-dimensional diabatic potential energy surfaces W_{11} (blue wavefunctions) and W_{22} (red wavefunctions) for the phenoxyl-phenol system. The top figure corresponds to the lowest-energy proton vibrational state.

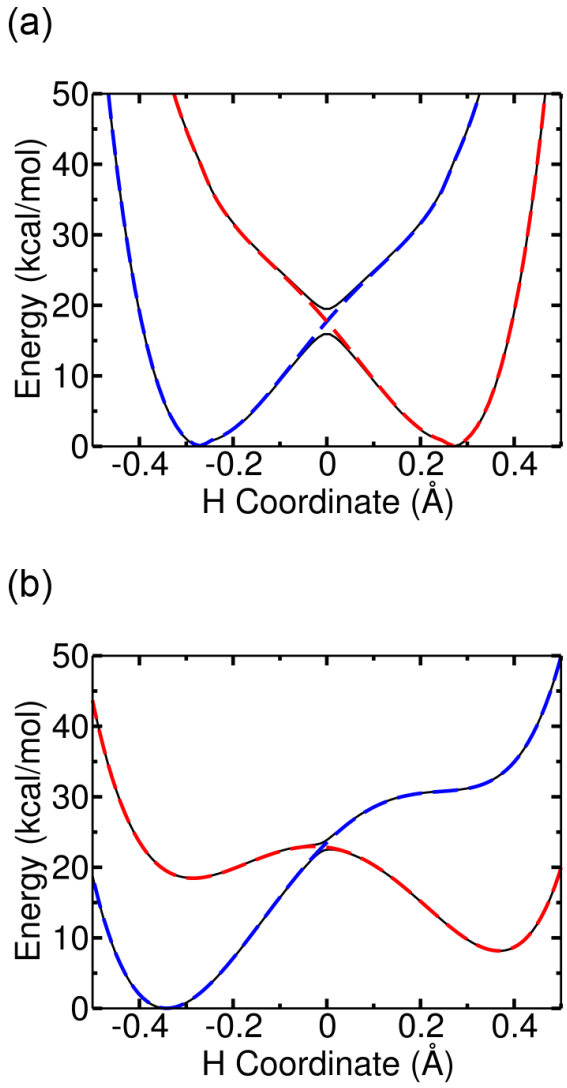

Figure 3.9: Electronically adiabatic and diabatic potential energy curves for (a) the phenoxyl-phenol system and (b) the amidinium-carboxylate system as functions of the hydrogen coordinate calculated using the GMH method. The solid black curves are the ground and first excited adiabatic state energies, and the dashed blue and red curves are the diabatic electronic energies.

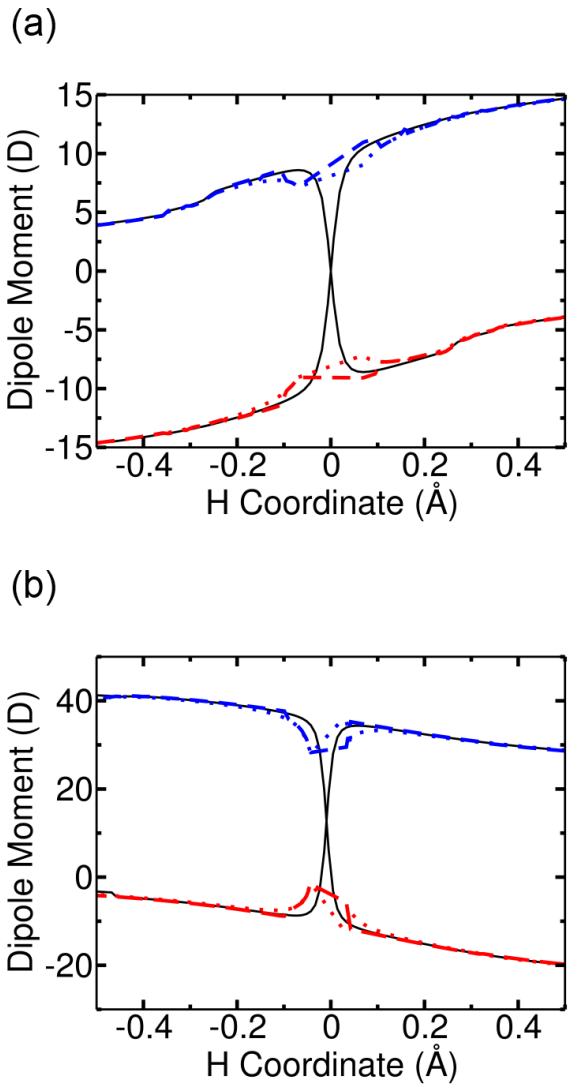

Figure 3.10: Component of the dipole moment vector along the hydrogen donor-acceptor axis for the ground and first excited adiabatic electronic states (solid black curves) and the diabatic electronic states ξ_1 (blue) and ξ_2 (red) for the (a) phenoxyl-phenol system and (b) amidinium-carboxylate system calculated using the GMH method (dotted) and the adiabatic-to-diabatic transformation method described in Section 3.2A (dashed). A positive (negative) dipole moment indicates a dipole moment vector pointing toward the acceptor (donor). The origin was chosen to be $r_p = 0$ for the amidinium-carboxylate system.

References

(1) R. I. Cukier and D. G. Nocera, *Annual Review of Physical Chemistry* **49**, 337 (1998).
(2) S. Hammes-Schiffer and A. V. Soudackov, *Journal of Physical Chemistry B* **112**, 14108 (2008).
(3) S. Hammes-Schiffer and A. A. Stuchebrukhov, *Chemical Reviews* **110**, 6939 (2010).
(4) M. H. V. Huynh and T. J. Meyer, *Chemical Reviews* **107**, 5004 (2007).
(5) J. M. Mayer, *Annual Review of Physical Chemistry* **55**, 363 (2004).
(6) S. Hammes-Schiffer, *Accounts of Chemical Research* **34**, 273 (2001).
(7) A. Soudackov and S. Hammes-Schiffer, *Journal of Chemical Physics* **113**, 2385 (2000).
(8) A. Soudackov, E. Hatcher, and S. Hammes-Schiffer, *Journal of Chemical Physics* **122**, 014505 (2005).
(9) G. J. Atchity and K. Ruedenberg, *Theoretical Chemistry Accounts* **97**, 47 (1997).
(10) M. Baer, *Chemical Physics Letters* **35**, 112 (1975).
(11) M. Baer, *Molecular Physics* **40**, 1011 (1980).
(12) M. Baer, *Physics Reports* **358**, 75 (2002).
(13) R. J. Cave and M. D. Newton, *Chemical Physics Letters* **249**, 15 (1996).
(14) R. J. Cave and M. D. Newton, *Journal of Chemical Physics* **106**, 9213 (1997).
(15) A. Cembran, L. C. Song, Y. R. Mo, and J. L. Gao, *Journal of Chemical Theory and Computation* **5**, 2702 (2009).
(16) H. Nakamura and D. G. Truhlar, *Journal of Chemical Physics* **115**, 10353 (2001).
(17) H. Nakamura and D. G. Truhlar, *Journal of Chemical Physics* **117**, 5576 (2002).
(18) M. D. Newton, *Chemical Reviews* **91**, 767 (1991).
(19) T. Pacher, L. S. Cederbaum, and H. Koppel, *Journal of Chemical Physics* **89**, 7367 (1988).
(20) T. Pacher, L. S. Cederbaum, and H. Koppel, *Advances in Chemical Physics* **84**, 293 (1993).
(21) B. N. Papas, M. S. Schuurman, and D. R. Yarkony, *Journal of Chemical Physics* **129**, 124104 (2008).
(22) K. Ruedenberg and G. J. Atchity, *Journal of Chemical Physics* **99**, 3799 (1993).
(23) A. Sirjoosingh and S. Hammes-Schiffer, *Journal of Physical Chemistry A* **115**, 2367 (2011).
(24) J. E. Subotnik, R. J. Cave, R. P. Steele, and N. Shenvi, *Journal of Chemical Physics* **130**, 234102 (2009).
(25) J. E. Subotnik, S. Yeganeh, R. J. Cave, and M. A. Ratner, *Journal of Chemical Physics* **129**, 244101 (2008).
(26) A. Thiel and H. Koppel, *Journal of Chemical Physics* **110**, 9371 (1999).
(27) T. Van Voorhis, T. Kowalczyk, B. Kaduk, L. P. Wang, C. L. Cheng, and Q. Wu, *Annual Review of Physical Chemistry* **61**, 149 (2010).
(28) Q. Wu and T. Van Voorhis, *Journal of Chemical Physics* **125**, 164105 (2006).
(29) Q. Wu and T. Van Voorhis, *Journal of Chemical Theory and Computation* **2**, 765 (2006).
(30) X. L. Zhu and D. R. Yarkony, *Journal of Chemical Physics* **132**, 104101 (2010).
(31) S. Hammes-Schiffer, *ChemPhysChem* **3**, 33 (2002).
(32) J. H. Skone, A. V. Soudackov, and S. Hammes-Schiffer, *Journal of the American Chemical Society* **128**, 16655 (2006).

(33) J. P. Kirby, J. A. Roberts, and D. G. Nocera, *Journal of the American Chemical Society* **119**, 9230 (1997).
(34) J. A. Roberts, J. P. Kirby, and D. G. Nocera, *Journal of the American Chemical Society* **117**, 8051 (1995).
(35) A. Soudackov and S. Hammes-Schiffer, *Journal of the American Chemical Society* **121**, 10598 (1999).
(36) Y. Georgievskii and A. A. Stuchebrukhov, *Journal of Chemical Physics* **113**, 10438 (2000).
(37) A. D. Becke, *Journal of Chemical Physics* **98**, 5648 (1993).
(38) C. T. Lee, W. T. Yang, and R. G. Parr, *Physical Review B* **37**, 785 (1988).
(39) P. C. Hariharan and J. A. Pople, *Theoretica Chimica Acta* **28**, 213 (1973).
(40) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian 09, Revision B.01 (Wallingford CT, 2009).
(41) M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, *Journal of Computational Chemistry* **14**, 1347 (1993).
(42) C. C. Marston and G. G. Balint-Kurti, *Journal of Chemical Physics* **91**, 3571 (1989).
(43) S. P. Webb and S. Hammes-Schiffer, *Journal of Chemical Physics* **113**, 5214 (2000).
(44) M. L. Connolly, *Journal of Applied Crystallography* **16**, 548 (1983).
(45) U. C. Singh and P. A. Kollman, *Journal of Computational Chemistry* **5**, 129 (1984).

Chapter 4

Derivation of an Electron-Proton Correlation Functional for Multicomponent Density Functional Theory within the Nuclear-Electronic Orbital Approach[‡]

Conventional density functional theory (DFT) relies on the Born-Oppenheimer separation of electrons and nuclei, and typically the nuclei move classically on adiabatic electronic surfaces. Nuclear quantum effects have been shown to be important for a broad range of systems, particularly those involving hydrogen bonding and hydrogen transfer.[1-3] In some cases, such as proton-coupled electron transfer reactions, nonadiabatic effects between the electrons and transferring protons have been shown to be significant.[4-6] Multicomponent DFT is a computationally practical method for incorporating these types of nuclear quantum effects into electronic structure calculations.[7-13] In the implementation of multicomponent DFT within the framework of the nuclear-electronic orbital (NEO) approach,[11,12,14] electrons and selected hydrogen nuclei are treated quantum mechanically without the Born-Oppeneheimer approximation. This NEO-DFT approach is designed for systems in which at least two nuclei are treated classically, eliminating difficulties associated with translations and rotations. Moreover, typically only a relatively small number of hydrogen nuclei, such as those involved in

[‡]Reproduced with permission from:
A. Sirjoosingh, M. V. Pak, and S. Hammes-Schiffer, "Derivation of an Electron-Proton Correlation Functional for Multicomponent Density Functional Theory within the Nuclear-Electronic Orbital Approach," *J. Chem. Theory Comput.* **7**, 2689-2693 (2011).

hydrogen bonding or hydrogen transfer, are treated quantum mechanically.

A major challenge of this approach is the development of electron-proton density functionals that accurately describe electron-proton correlation, which is highly significant because of the attractive interaction between the electron and proton and the disparity in the masses.[15,16] Previously we devised a strategy for the development of electron-proton density functionals using the electron-proton pair density from an explicitly correlated nuclear-electronic wavefunction.[11] The initial electron-proton density functional developed with this strategy required the neglect of a large number of terms in the explicitly correlated electron-proton pair density. The objective of the present work is to derive an electron-proton density functional with a different ansatz for the explicitly correlated nuclear-electronic wavefunction, thereby enabling us to retain all of the terms in the electron-proton pair density. The resulting electron-proton functional has a similar form as the previous functional but is expected to be more reliable in terms of scaling with respect to the number of electrons and quantum protons. The application of this new functional to model systems illustrates that it provides accurate hydrogen nuclear densities. Moreover, the form of this functional is computationally practical for larger molecular systems.

We consider a multicomponent system comprised of N_e electrons and N_p protons that are treated quantum mechanically in a field of N_c fixed classical nuclei. Within the framework of multicomponent DFT, the ground state energy is the minimum of the energy functional

$$E\left[\rho^e,\rho^p\right]=\int d\mathbf{r}_1^e\rho^e\left(\mathbf{r}_1^e\right)v\left(\mathbf{r}_1^e\right)-\int d\mathbf{r}_1^p\rho^p\left(\mathbf{r}_1^p\right)v\left(\mathbf{r}_1^p\right)+F\left[\rho^e,\rho^p\right] \tag{4.1}$$

subject to the constraints $\int d\mathbf{r}_1^e\rho^e\left(\mathbf{r}_1^e\right)=N_e$ and $\int d\mathbf{r}_1^p\rho^p\left(\mathbf{r}_1^p\right)=N_p$. Here $\mathbf{r}^e$ and $\mathbf{r}^p$ denote the collective spatial coordinates of the electrons and quantum protons, respectively, $\rho^e\left(\mathbf{r}_1^e\right)$ and

$\rho^p\left(\mathbf{r}_1^p\right)$ denote the one-particle electron and proton densities, respectively, and $v(\mathbf{r}_1)$ is the Coulomb interaction between the electron or proton and the classical nuclei, as defined in Ref. 12. Analogous to electronic DFT, we define a noninteracting reference system in which all quantum particles (i.e., electrons and quantum protons) do not interact with each other. The ground state nuclear-electronic wavefunction of this noninteracting reference system is given by the product of electronic and nuclear Slater determinants.

Following the Kohn-Sham procedure,[17,18] the universal functional $F\left[\rho^e,\rho^p\right]$ for the interacting system can be expressed as[11,12]

$$\begin{aligned} F\left[\rho^e,\rho^p\right] = T_s\left[\rho^e,\rho^p\right] + J_{ep}\left[\rho^e,\rho^p\right] + E_{epc}\left[\rho^e,\rho^p\right] \\ + J_{ee}\left[\rho^e\right] + E_{exc}\left[\rho^e\right] + J_{pp}\left[\rho^p\right] + E_{pxc}\left[\rho^p\right], \end{aligned} \tag{4.2}$$

where $T_s\left[\rho^e,\rho^p\right]$ is the total kinetic energy for the noninteracting system. The classical parts of the electron-proton and electron-electron Coulomb interactions are given by

$$J_{ep}\left[\rho^e,\rho^p\right] = -\iint \mathrm{d}\mathbf{r}_1^e \mathrm{d}\mathbf{r}_1^p \, \frac{\rho^e\left(\mathbf{r}_1^e\right)\rho^p\left(\mathbf{r}_1^p\right)}{\left|\mathbf{r}_1^e - \mathbf{r}_1^p\right|} \tag{4.3}$$

and

$$J_{ee}\left[\rho^e\right] = \frac{1}{2}\iint \mathrm{d}\mathbf{r}_1^e \mathrm{d}\mathbf{r}_2^e \, \frac{\rho^e\left(\mathbf{r}_1^e\right)\rho^e\left(\mathbf{r}_2^e\right)}{\left|\mathbf{r}_1^e - \mathbf{r}_2^e\right|}, \tag{4.4}$$

and the proton-proton Coulomb interaction $J_{pp}\left[\rho^p\right]$ is defined analogously.

The terms $E_{epc}\left[\rho^e,\rho^p\right]$, $E_{exc}\left[\rho^e\right]$, and $E_{pxc}\left[\rho^p\right]$ are the electron-proton correlation functional, the electron exchange-correlation functional, and the proton exchange-correlation functional, respectively. In this formulation,[11,12] the definition of the electron exchange-

correlation functional, $E_{exc}\left[\rho^e\right]$, is consistent with that from standard electronic DFT.[17-19] Thus, the traditional, well-established electron exchange-correlation functionals can be used, although these electronic functionals have been parametrized without the inclusion of nuclear quantum effects and electron-proton correlation. Furthermore, the contribution from the proton exchange-correlation functional, $E_{pxc}\left[\rho^p\right]$, is assumed to be negligible due to the localized nature of protons in typical molecular systems with only selected hydrogen nuclei treated quantum mechanically. For systems with multiple quantum nuclei, the quantum protons may be treated with a generalized Hartree-Fock approach, in which each proton can occupy a different localized spatial orbital, and the proton exchange-correlation functional may be chosen to be the diagonal proton exchange interaction terms to eliminate the self-interaction terms.[20] The present paper focuses on the development of a suitable electron-proton correlation functional, $E_{epc}\left[\rho^e,\rho^p\right]$.

Following the strategy devised in Ref. 11, we define the electron-proton correlation functional in terms of the electron-proton pair density, $\rho^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)$, as

$$E_{epc}\left[\rho^e,\rho^p\right]=-\iint \mathrm{d}\mathbf{r}_1^e\mathrm{d}\mathbf{r}_1^p\,\frac{\rho^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)}{\left|\mathbf{r}_1^e-\mathbf{r}_1^p\right|}-J_{ep}\left[\rho^e,\rho^p\right]. \tag{4.5}$$

In Ref. 11, the electron-proton pair density was obtained from an explicitly correlated nuclear-electronic wavefunction defined as $\Psi_{\mathrm{gem}}=(1+G)\Phi^e\Phi^p$,[15] where Φ^e and Φ^p are electronic and nuclear Slater determinants, respectively, and

$$G\left(\mathbf{r}^e,\mathbf{r}^p\right)=\sum_{i=1}^{N_e}\sum_{i'=1}^{N_p}g\left(\mathbf{r}_i^e,\mathbf{r}_{i'}^p\right), \tag{4.6}$$

$$g\left(\mathbf{r}_i^e,\mathbf{r}_{i'}^p\right)=\sum_{k=1}^{N_{\mathrm{gem}}}b_k e^{-\gamma_k\left|\mathbf{r}_i^e-\mathbf{r}_{i'}^p\right|^2}\quad . \tag{4.7}$$

Here N_{gem} is the number of Gaussian type geminal functions used in the expansion, and b_k and γ_k are parameters that define these functions. Gaussian type geminal functions are used to ensure the tractable calculation of integrals over Gaussian basis functions.

In the present paper, we determine the electron-proton pair density from the alternative explicitly correlated nuclear-electronic wavefunction defined as[21]

$$\Psi_{\text{gem}} = \sqrt{1+G}\,\Phi^e\Phi^p. \tag{4.8}$$

The significant advantage of this alternative nuclear-electronic wavefunction ansatz over the previous ansatz is that all terms quadratic in the geminal functions are eliminated from the electron-proton pair density. Note that this alternative ansatz retains the important characteristics of a mixed nuclear-electronic wavefunction: it is antisymmetric with respect to exchange of electrons or quantum protons, approaches the Hartree-Fock wavefunction at large electron-proton distances, and has the numerical flexibility to describe the correct linear behavior at small electron-proton distances. As a result of this alternative ansatz, in conjunction with a physically reasonable approximation for the two-particle electron and proton densities, all terms in the electron-proton pair density can be retained in the present treatment, compared to the neglect of 24 out of 26 terms in the electron-proton pair density in the previous treatment.[11,22] The remainder of this Letter presents the derivation of the new electron-proton density functional and an initial application to a model system.

Prior to the derivation, we define the geminal reduced densities, which are associated with the geminal wavefunction, and the auxiliary reduced densities, which are associated with the Slater determinants. The geminal one-particle and two-particle electron densities are defined as

$$\rho_1^e\left(\mathbf{r}_1^e\right)=\frac{N_e}{\left\langle\Psi_{\text{gem}}\middle|\Psi_{\text{gem}}\right\rangle}\left\langle\Psi_{\text{gem}}\middle|\Psi_{\text{gem}}\right\rangle_{-e_1} \tag{4.9}$$

$$\rho_2^e\left(\mathbf{r}_1^e,\mathbf{r}_2^e\right)=\frac{N_e(N_e-1)}{2\left\langle\Psi_{\text{gem}}\middle|\Psi_{\text{gem}}\right\rangle}\left\langle\Psi_{\text{gem}}\middle|\Psi_{\text{gem}}\right\rangle_{-e_1e_2} \tag{4.10}$$

and the one-particle and two-particle proton densities are defined analogously. The geminal electron-proton pair density is defined as

$$\rho^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)=\frac{N_eN_p}{\left\langle\Psi_{\text{gem}}\middle|\Psi_{\text{gem}}\right\rangle}\left\langle\Psi_{\text{gem}}\middle|\Psi_{\text{gem}}\right\rangle_{-e_1p_1}. \tag{4.11}$$

In these expressions and those that follow, angular brackets without subscripts indicate integration over all coordinates, angular brackets with subscripts indicate the spatial coordinates of integration (i.e., $\langle\cdots\rangle_{e_1}$ indicates integration over $\mathbf{r}_1^e$), and angular brackets with subscripts preceded by a minus sign denote integration over all coordinates except the specified spatial coordinate(s) (i.e., $\langle\cdots\rangle_{-e_1}$ indicates integration over all coordinates except $\mathbf{r}_1^e$). The auxiliary one-particle and two-particle electron densities, $\tilde{\rho}_1^e\left(\mathbf{r}_1^e\right)$ and $\tilde{\rho}_2^e\left(\mathbf{r}_1^e,\mathbf{r}_2^e\right)$, are defined by substituting Φ^e for Ψ_{gem} in Eqs. (4.9) and (4.10), and the auxiliary proton densities are defined analogously with Φ^p. Note that the Slater determinants, Φ^e and Φ^p, are normalized because they are constructed with orthonormal spin orbitals, but the geminal wavefunction, Ψ_{gem}, is not normalized due to the geminal factor in Eq. (4.8).

The geminal electron-proton pair density corresponding to the ansatz given in Eq. (4.8) can be expressed in terms of the auxiliary densities as follows:

$$\rho^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)=\frac{1}{1+\left\langle\tilde{\rho}_1^e\left(\mathbf{r}_1^e\right)\tilde{\rho}_1^p\left(\mathbf{r}_1^p\right)g\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)\right\rangle_{e_1p_1}}\Big[\tilde{\rho}_1^e\left(\mathbf{r}_1^e\right)\tilde{\rho}_1^p\left(\mathbf{r}_1^p\right)\left\{1+g\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)\right\}$$
$$+2\tilde{\rho}_1^p\left(\mathbf{r}_1^p\right)\left\langle\tilde{\rho}_2^e\left(\mathbf{r}_1^e,\mathbf{r}_2^e\right)g\left(\mathbf{r}_2^e,\mathbf{r}_1^p\right)\right\rangle_{e_2}+2\tilde{\rho}_1^e\left(\mathbf{r}_1^e\right)\left\langle\tilde{\rho}_2^p\left(\mathbf{r}_1^p,\mathbf{r}_2^p\right)g\left(\mathbf{r}_1^e,\mathbf{r}_2^p\right)\right\rangle_{p_2} \quad (4.12)$$
$$+4\left\langle\tilde{\rho}_2^e\left(\mathbf{r}_1^e,\mathbf{r}_2^e\right)\tilde{\rho}_2^p\left(\mathbf{r}_1^p,\mathbf{r}_2^p\right)g\left(\mathbf{r}_2^e,\mathbf{r}_2^p\right)\right\rangle_{e_2p_2}\Big].$$

For comparison, the electron-proton pair density corresponding to the previous wavefunction ansatz, with the geminal factor of (1+*G*) rather than $\sqrt{1+G}$, was comprised of 26 terms, as given by Eq. (32) in Ref. 22. Thus, the present electron-proton pair density is much simpler. In the previous derivation,[11] the geminal electron-proton pair density was truncated in a manner that eliminated all terms of order g^2 and all terms that included densities other than one-particle densities. As a result, the electron-proton pair density included only the first two terms, $\tilde{\rho}_1^e\tilde{\rho}_1^p(1+g)$, in Eq. (4.12). This truncation required the renormalization of the electron-proton pair density so that $\left\langle\rho^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)\right\rangle_{e_1p_1}=N_eN_p$, leading to an additional factor of $\left(N_eN_p\right)^{-1}$ in the second term of the denominator. In the present approach, the electron-proton pair density in Eq. (4.12) is already normalized properly because no terms have been eliminated. Note that the wavefunction ansatz with the geminal factor of $\sqrt{1+G}$ does not lead to any terms of order g^2 in the electron-proton pair density, and the terms including two-particle densities have been retained in Eq. (4.12).

To develop an effective electron-proton density functional, the electron-proton pair density given in Eq. (4.12) should depend on only one-particle densities. For this purpose, we assume that the auxiliary two-particle electron density can be approximated as

$$\tilde{\rho}_2^e\left(\mathbf{r}_1^e,\mathbf{r}_2^e\right)\approx\frac{1}{2}\frac{N_e-1}{N_e}\tilde{\rho}_1^e\left(\mathbf{r}_1^e\right)\tilde{\rho}_1^e\left(\mathbf{r}_2^e\right). \quad (4.13)$$

This independent particle approximation is based on the assumption that the direct electron-

electron exchange contributions included in the electron exchange-correlation functional, as defined in Eq. (4.2), are significantly greater than the indirect electron-electron exchange contributions in the electron-proton correlation functional. As a result, the indirect electron-electron exchange effects arising from the dependence of the electron-proton pair density on the two-particle electron density are neglected. We invoke the analogous approximation for the auxiliary two-particle proton density. These two-particle density terms were completely neglected in the previous treatment.[11]

Substituting these approximate auxiliary two-particle electron and proton densities into Eq. (4.12) leads to

$$\rho^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)=\frac{\tilde{\rho}^e\tilde{\rho}^p}{1+\left\langle\tilde{\rho}^e\tilde{\rho}^p g\right\rangle_{ep}}\left[1+g+\frac{N_e-1}{N_e}\left\langle\tilde{\rho}^e g\right\rangle_e+\frac{N_p-1}{N_p}\left\langle\tilde{\rho}^p g\right\rangle_p+\frac{\left(N_e-1\right)\left(N_p-1\right)}{N_eN_p}\left\langle\tilde{\rho}^e\tilde{\rho}^p g\right\rangle_{ep}\right]. \tag{4.14}$$

For notational convenience, we have dropped the dependence of the reduced densities on the coordinates, defined $g\equiv g(\mathbf{r}_1^e,\mathbf{r}_1^p)$, dropped the subscript on the one-particle densities, and simplified the subscripts on the brackets to denote the electron and/or proton spatial coordinates in the integrand. The analogous procedure for the geminal one-particle electron and proton densities leads to:

$$\rho^e\left(\mathbf{r}_1^e\right)=\frac{\tilde{\rho}^e}{1+\left\langle\tilde{\rho}^e\tilde{\rho}^p g\right\rangle_{ep}}\left(1+\left\langle\tilde{\rho}^p g\right\rangle_p+\frac{N_e-1}{N_e}\left\langle\tilde{\rho}^e\tilde{\rho}^p g\right\rangle_{ep}\right) \tag{4.15}$$

$$\rho^p\left(\mathbf{r}_1^p\right)=\frac{\tilde{\rho}^p}{1+\left\langle\tilde{\rho}^e\tilde{\rho}^p g\right\rangle_{ep}}\left(1+\left\langle\tilde{\rho}^e g\right\rangle_e+\frac{N_p-1}{N_p}\left\langle\tilde{\rho}^e\tilde{\rho}^p g\right\rangle_{ep}\right). \tag{4.16}$$

Note that these densities satisfy the sum rules, $\rho^e\left(\mathbf{r}_1^e\right)=N_p^{-1}\left\langle\rho^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)\right\rangle_p$ and

$\rho^p\left(\mathbf{r}_1^p\right) = N_e^{-1}\left\langle \rho^{ep}\left(\mathbf{r}_1^e, \mathbf{r}_1^p\right)\right\rangle_e$.

The next step is to express the electron-proton pair density in Eq. (4.14) in terms of the one-particle densities given in Eqs. (4.15) and (4.16) by eliminating the auxiliary densities. To achieve this goal, expressions for the auxiliary one-particle densities in terms of the geminal one-particle densities must be determined. In principle, Eqs. (4.15) and (4.16) could be inverted to determine these expressions, but the exact analytical solution is not known. Instead, we follow the approximate procedure of Ref. 11 and replace $\tilde{\rho}^e\left(\mathbf{r}_1^e\right)$ with $\rho^e\left(\mathbf{r}_1^e\right)$ and $\tilde{\rho}^p\left(\mathbf{r}_1^p\right)$ with $\rho^p\left(\mathbf{r}_1^p\right)$ whenever they are multiplied by the geminal factor *g* in Eqs. (4.14), (4.15), and (4.16). These substitutions lead to the following expressions for the approximate geminal densities:

$$\rho^{ep}\left(\mathbf{r}_1^e, \mathbf{r}_1^p\right) = \frac{1}{1+\left\langle \rho^e \rho^p g\right\rangle_{ep}}\left[\tilde{\rho}^e\tilde{\rho}^p + \rho^e\rho^p g + \frac{N_e-1}{N_e}\rho^e\rho^p\left\langle \rho^e g\right\rangle_e + \frac{N_p-1}{N_p}\rho^e\rho^p\left\langle \rho^p g\right\rangle_p + \frac{\left(N_e-1\right)\left(N_p-1\right)}{N_e N_p}\rho^e\rho^p\left\langle \rho^e\rho^p g\right\rangle_{ep}\right] \tag{4.17}$$

$$\rho^e\left(\mathbf{r}_1^e\right) = \frac{1}{1+\left\langle \rho^e\rho^p g\right\rangle_{ep}}\left[\tilde{\rho}^e + \rho^e\left\langle \rho^p g\right\rangle_p + \frac{N_e-1}{N_e}\rho^e\left\langle \rho^e\rho^p g\right\rangle_{ep}\right] \tag{4.18}$$

$$\rho^p\left(\mathbf{r}_1^p\right) = \frac{1}{1+\left\langle \rho^e\rho^p g\right\rangle_{ep}}\left[\tilde{\rho}^p + \rho^p\left\langle \rho^e g\right\rangle_e + \frac{N_p-1}{N_p}\rho^p\left\langle \rho^e\rho^p g\right\rangle_{ep}\right]. \tag{4.19}$$

Note that these reduced densities still satisfy the sum rules given above and retain the property that $\lim_{r_{ep}\to\infty}\rho^{ep} = \tilde{\rho}^e\tilde{\rho}^p = \rho^e\rho^p$.

Substituting the expressions for the auxiliary one-particle electron and proton densities obtained from Eqs. (4.18) and (4.19) into the electron-proton pair density given in Eq. (4.17) leads to the final expression for the approximate pair density in terms of the one-particle

densities:

$$\rho^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)=\rho^e\rho^p\left[1+\left\langle N_e^{-1}N_p^{-1}\rho^e\rho^p g\right\rangle_{ep}-\left\langle N_e^{-1}\rho^e g\right\rangle_e-\left\langle N_p^{-1}\rho^p g\right\rangle_p + \frac{g+\left\langle\rho^e g\right\rangle_e\left\langle\rho^p g\right\rangle_p}{1+\left\langle\rho^e\rho^p g\right\rangle_{ep}}\right]. \tag{4.20}$$

Eq. (4.20) represents an approximate pair density derived from the explicitly correlated nuclear-electronic wavefunction given in Eq. (4.8) and defines an electron-proton correlation functional when substituted into Eq. (4.5). This expression is identical to the previous expression[11] derived from a different explicitly correlated nuclear-electronic wavefunction except the last term, which differs by a factor of N_eN_p in the second term of both the numerator and denominator. These differences arise from the truncation of the electron-proton pair density and the subsequent renormalization in the previous treatment. Since the present derivation includes all terms of the electron-proton pair density, whereas the previous derivation neglected a large number of terms, we expect the present functional to be more reliable in terms of scaling with respect to the number of electrons and protons. The computational cost is identical for the two functionals.

We applied the NEO-DFT approach with this electron-proton density functional to the model system, $[He\text{-}H\text{-}He]^+$, and the isotopomers $[He\text{-}D\text{-}He]^+$ and $[He\text{-}T\text{-}He]^+$. The two helium nuclei were treated classically at a fixed distance, and the central nucleus and four electrons were treated quantum mechanically. We studied these systems with the cc-pVDZ and cc-pVTZ electronic basis sets,[23,24] where the electronic basis functions corresponding to the central nucleus were placed at the midpoint between the two helium nuclei. The nuclear basis set was comprised of a single 1s nuclear basis function placed at the midpoint between the two helium nuclei, and the exponent was optimized variationally during the NEO-DFT calculation. Two Gaussian type geminal functions were used with geminal parameters obtained variationally with

the wavefunction ansatz in Eq. (4.8) for a one-electron/one-proton model system.[21] To account for differences between the NEO-DFT and variational wavefunction approach, the geminal functions were scaled by a single constant factor to reproduce the hydrogen vibrational stretching frequency of $[He\text{-}H\text{-}He]^+$ with the cc-pVDZ electronic basis set. For all other systems and basis sets, these geminal parameters were fixed during the NEO-DFT calculations. Moreover, in this Letter, the electron exchange-correlation functional was chosen to be the Hartree-Fock exchange. Future studies will examine the effects of combining this electron-proton correlation functional with various electron exchange-correlation functionals. All calculations were performed with a modified version of the GAMESS program.[25]

The objective of this application is to provide evidence that this electron-proton functional can provide accurate hydrogen nuclear densities. The hydrogen vibrational stretching frequencies were determined from a Gaussian fit of the nuclear density along the He-He axis. These frequencies are compared to the corresponding splitting for the three-dimensional hydrogen vibrational states calculated with the Fourier grid Hamiltonian (FGH) method.[26] Since the nuclear basis set contains only a single 1s nuclear basis function, it is incapable of reproducing both the stretching and bending hydrogen vibrational frequencies. For these calculations, we determined the He-He distances at which the stretching and bending frequencies calculated with the FGH method are qualitatively similar. The resulting He-He distances for the cc-pVDZ and cc-pVTZ electronic basis sets were determined to be 1.955 Å and 1.945 Å, respectively. Future work will focus on studies with larger electronic and nuclear basis sets that will enable the calculation of bending as well as stretching frequencies.

The results of these calculations are provided in Table 4.1. The NEO-HF (Hartree-Fock) frequencies are much higher than the NEO-DFT frequencies, which are in qualitative agreement

with the FGH frequencies. These values illustrate the importance of electron-proton correlation. In addition, these results indicate that the geminal parameters are reasonably transferable to larger electronic basis sets and to other isotopes of hydrogen for this model system. From a physical perspective, the geminal parameters are expected to be transferable because these terms are significant only at small electron-proton distances and should be relatively independent of the external chemical environment. Future work will focus on optimizing the geminal parameters for applications to a wide range of chemical systems. Qualitatively similar results are obtained for this model system with the previously derived electron-proton functional[11] using different geminal parameters, as given in Appendix C. Note that this previous functional may be derived from the ansatz given in Eq. (4.8) if the two-particle densities are neglected. The two electron-proton functionals will exhibit different scaling behavior with respect to the number of electrons and protons, however, and the present functional is expected to be more reliable because fewer approximations were invoked in the derivation. Investigation of these scaling properties will require the study of systems with a larger number of quantum particles and is a direction for future research.

In this Letter, we derived an electron-proton density functional for use in multicomponent DFT calculations, where electrons and selected hydrogen nuclei are treated quantum mechanically. This functional was derived directly from the electron-proton pair density associated with a recently proposed ansatz for the explicitly correlated nuclear-electronic wavefunction. The advantage of this functional over the previously derived functional is that the new wavefunction ansatz, combined with the independent particle approximation for the auxiliary two-particle electron and proton densities, enabled us to retain all of the terms in the electron-proton pair density, whereas the previous derivation neglected a large number of terms.

Thus, the present functional is based on a more rigorous derivation and therefore may be more robust, although further studies are required to assess both functionals. In particular, future work will focus on the further development of these types of electron-proton functionals with larger electronic and nuclear basis sets in conjunction with standard electron exchange-correlation functionals.

Isotope	cc-pVDZ			cc-pVTZ		
	NEO-HF	NEO-DFT	FGH	NEO-HF	NEO-DFT	FGH
H	3098	1191	1191	3122	1103	1111
D	2284	820	801	2330	782	740
T	1903	660	633	1954	646	581

Table 4.1: Vibrational frequencies in cm^{-1} corresponding to the hydrogen vibrational stretching motion calculated with the NEO-HF, NEO-DFT, and FGH methods for the $[He\text{-}X\text{-}He]^+$ systems with X = H, D, or T. The cc-pVDZ or cc-pVTZ electronic basis set was used as indicated. The NEO-HF and NEO-DFT calculations were performed using a single 1*s* nuclear basis set with a variationally optimized exponent, and the NEO-DFT calculations were performed using two geminals with parameters $(b_1, \gamma_1) = (0.3969, 0.34)$ and $(b_2, \gamma_2) = (0.8912, 2.47)$ for the electron-proton functional presented in this paper.

References

(1) M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello, *Science* **275**, 817 (1997).
(2) S. Raugei and M. L. Klein, *Journal of the American Chemical Society* **125**, 8992 (2003).
(3) Y. Cha, C. J. Murray, and J. P. Klinman, *Science* **243**, 1325 (1989).
(4) S. Hammes-Schiffer and A. A. Stuchebrukhov, *Chemical Reviews* **110**, 6939 (2010).
(5) A. Sirjoosingh and S. Hammes-Schiffer, *Journal of Physical Chemistry A* **115**, 2367 (2011).
(6) Y. Georgievskii and A. A. Stuchebrukhov, *Journal of Chemical Physics* **113**, 10438 (2000).
(7) J. F. Capitani, R. F. Nalewajski, and R. G. Parr, *Journal of Chemical Physics* **76**, 568 (1982).
(8) N. Gidopoulos, *Physical Review B* **57**, 2146 (1998).
(9) T. Kreibich and E. K. U. Gross, *Physical Review Letters* **86**, 2984 (2001).
(10) T. Kreibich, R. van Leeuwen, and E. K. U. Gross, *Physical Review A* **78**, 022501 (2008).
(11) A. Chakraborty, M. V. Pak, and S. Hammes-Schiffer, *Physical Review Letters* **101**, 153001 (2008).
(12) A. Chakraborty, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Physics* **131**, 124115 (2009).
(13) V. Krishna, *Physical Review Letters* **102**, 053002 (2009).
(14) S. P. Webb, T. Iordanov, and S. Hammes-Schiffer, *Journal of Chemical Physics* **117**, 4106 (2002).
(15) C. Swalina, M. V. Pak, A. Chakraborty, and S. Hammes-Schiffer, *Journal of Physical Chemistry A* **110**, 9983 (2006).
(16) A. Chakraborty, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Physics* **129**, 014101 (2008).
(17) W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).
(18) R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules*. (Oxford University Press, New York, 1989).
(19) M. Ernzerhof, J. P. Perdew, and K. Burke, in *Density Functional Theory*, edited by R. Nalewajski (Springer-Verlag, Berlin, 1996).
(20) B. Auer and S. Hammes-Schiffer, *Journal of Chemical Physics* **132**, 084110 (2010).
(21) C. Ko, M. V. Pak, C. Swalina, and S. Hammes-Schiffer, *Journal of Chemical Physics* **135**, 054106 (2011).
(22) A. Chakraborty and S. Hammes-Schiffer, *Journal of Chemical Physics* **129**, 204101 (2008).
(23) T. H. Dunning, Jr., *Journal of Chemical Physics* **90**, 1007 (1989).
(24) D. E. Woon and T. H. Dunning, Jr., *Journal of Chemical Physics* **100**, 2975 (1994).
(25) M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, *Journal of Computational Chemistry* **14**, 1347 (1993).
(26) S. P. Webb and S. Hammes-Schiffer, *Journal of Chemical Physics* **113**, 5214 (2000).

Chapter 5

Multicomponent Density Functional Theory Study of the Interplay between Electron-Electron and Electron-Proton Correlation‡

5.1. Introduction

Nuclear quantum effects play an important role in a wide variety of chemical systems, particularly those involving hydrogen bonding and hydrogen transfer.[1-3] In many cases, such as proton-coupled electron transfer reactions,[4,5] nonadiabatic effects between electrons and nuclei are significant. To accurately model these systems, methods that do not invoke the Born-Oppenheimer approximation must be developed. A promising method of this type is the nuclear-electronic orbital (NEO) approach, in which electrons and select protons are treated quantum mechanically on equal footing.[6] Electron-proton correlation has been found to be highly significant,[7-9] leading to the development of the explicitly correlated methods denoted NEO-XCHF[10,11] and NEO-XCHF2.[12] Although these explicitly correlated wavefunction methods are accurate for small molecular systems, currently they are computationally prohibitive for larger systems.

Conventional electronic density functional theory (DFT) methods have been successfully applied to relatively large molecular systems. Typical implementations of electronic DFT

‡Reproduced with permission from:
A. Sirjoosingh, M. V. Pak, and S. Hammes-Schiffer, "Multicomponent density functional theory study of the interplay between electron-electron and electron-proton correlation," *J. Chem. Phys.* **136**, 174114 (2012).

methods invoke the Born-Oppenheimer approximation with all nuclei treated classically and thus are not applicable to systems in which nuclear quantum effects and nonadiabatic effects between the electrons and nuclei are important. Multicomponent DFT approaches[13-20] have been developed to treat different types of particles quantum mechanically within the framework of DFT. In the NEO-DFT approach,[17,19-21] electrons and select protons are treated quantum mechanically without invoking the Born-Oppenheimer approximation. Electron-proton correlation effects are included directly with an electron-proton correlation functional.[17,20] A significant advantage of the NEO-DFT method over the explicitly correlated wavefunction methods is that electron-electron and electron-proton correlation are treated consistently. In addition, the NEO-DFT approach is more computationally practical for larger molecular systems.

Previously we developed electron-proton correlation functionals for use in the NEO-DFT approach from the electron-proton pair densities obtained from explicitly correlated wavefunctions.[17,20] In particular, the NEO-XCHF pair density was used to derive the EPC1 functional,[17] and the NEO-XCHF2 pair density was used to derive the EPC2 functional.[20] The derivations of both of these functionals relied on the assumption that the kinetic energy contribution arising from electron-proton correlation is negligible. In the present paper, we demonstrate how this kinetic energy contribution can be included in the electron-proton functional using the adiabatic connection formula in multicomponent DFT.[19] We apply this strategy to the EPC2 functional and thus derive the EPC2-KE functional, which includes the kinetic energy contribution arising from electron-proton correlation.

The EPC1 and EPC2 electron-proton correlation functionals were previously implemented and tested with Hartree-Fock exchange as the electronic exchange-correlation

functional.[17,20] In the present paper, we assess the performance of the three electron-proton correlation functionals discussed above (i.e., EPC1, EPC2 and EPC2-KE) with three different electron exchange-correlation functionals, namely B3LYP, BLYP and PBE.[22-24] Our calculations demonstrate that electron-proton and electron-electron correlation are predominantly uncoupled for the model systems studied. The consequence of this behavior is that electron-proton correlation functionals and electronic exchange-correlation functionals can be developed independently and subsequently combined without the necessity for re-parameterization of either type of functional. This characteristic of the NEO-DFT approach is advantageous for the study of diverse molecular systems for which well-established electronic exchange-correlation functionals have been shown to perform well.

An outline of the paper is as follows. In Section 5.2, we summarize the important aspects of multicomponent DFT and the development of the electron-proton correlation functionals used in this study. A detailed derivation of the new EPC2-KE functional is provided in Appendix D. Section 5.3 presents the application of the NEO-DFT method to model systems using these electron-proton correlation functionals combined with three different electronic exchange-correlation functionals. The conclusions and future directions are discussed in Section 5.4.

5.2. Theory

A. Multicomponent density functional theory

In this subsection, we present a formulation of multicomponent DFT for a system comprised of N_e electrons, N_p protons, and N_c fixed classical nuclei. The Hamiltonian in atomic units for this system is given by

$$H = -\frac{1}{2}\sum_{i=1}^{N_e}\nabla_i^2 - \frac{1}{2m_p}\sum_{i'=1}^{N_p}\nabla_{i'}^2 + \sum_{i=1}^{N_e} v\left(\mathbf{r}_i^e\right) - \sum_{i'=1}^{N_p} v\left(\mathbf{r}_{i'}^p\right) + \sum_{i=1}^{N_e}\sum_{j>i}^{N_e}\frac{1}{\left|\mathbf{r}_i^e - \mathbf{r}_j^e\right|} + \sum_{i'=1}^{N_p}\sum_{j'>i}^{N_p}\frac{1}{\left|\mathbf{r}_{i'}^p - \mathbf{r}_{j'}^p\right|} - \sum_{i=1}^{N_e}\sum_{i'=1}^{N_p}\frac{1}{\left|\mathbf{r}_i^e - \mathbf{r}_{i'}^p\right|} \quad , \tag{5.1}$$

where $\mathbf{r}^e$ and $\mathbf{r}^p$ denote the collective spatial coordinates of the electrons and quantum protons, respectively, m_p is the mass of the proton, and $v\left(\mathbf{r}_i\right)$ is the interaction between an electron or quantum proton with the classical nuclei. This interaction is defined as

$$v\left(\mathbf{r}_i\right) = -\sum_{A=1}^{N_c}\frac{Z_A}{\left|\mathbf{r}_i - \mathbf{r}_A^c\right|} \quad , \tag{5.2}$$

where $\mathbf{r}_A^c$ and Z_A are the coordinate and charge, respectively of the A^{th} classical nucleus. For clarity, we use primed and unprimed indices for quantum protons and electrons, respectively. Note that the term corresponding to repulsion between classical nuclei is omitted from Eq. (5.1) for simplicity.

Within the framework of the NEO-DFT approach, the ground state energy is the minimum of the functional

$$E\left[\rho^e, \rho^p\right] = \int \mathrm{d}\mathbf{r}_1^e \rho^e\left(\mathbf{r}_1^e\right) v\left(\mathbf{r}_1^e\right) - \int \mathrm{d}\mathbf{r}_1^p \rho^p\left(\mathbf{r}_1^p\right) v\left(\mathbf{r}_1^p\right) + F\left[\rho^e, \rho^p\right] \tag{5.3}$$

subject to the constraints $\int \mathrm{d}\mathbf{r}_1^e \rho^e\left(\mathbf{r}_1^e\right) = N_e$ and $\int \mathrm{d}\mathbf{r}_1^p \rho^p\left(\mathbf{r}_1^p\right) = N_p$ for the electron and proton densities, ρ^e and ρ^p, respectively. The exact universal functional can be defined using Levy's constrained search formulation[25] as

$$F\left[\rho^e, \rho^p\right] = \min_{\Psi \to \rho^e \rho^p} \left\langle \Psi \middle| T + V \middle| \Psi \right\rangle = \left\langle \Psi_{\min}^{\rho^e \rho^p} \middle| T + V \middle| \Psi_{\min}^{\rho^e \rho^p} \right\rangle \tag{5.4}$$

where T and V represent the total kinetic and potential energy operators, respectively. In this formulation, the search is over all wavefunctions that are antisymmetric with respect to exchange

of like particles (i.e., electrons or protons) and correspond to electronic and nuclear densities ρ^e and ρ^p, respectively. The minimizing wavefunction, which is assumed to exist, is denoted $\Psi_{\min}^{\rho^e \rho^p}$. For convenience, we partition the kinetic and potential energy operators as

$$T = T_{ee} + T_{pp} = \left(-\frac{1}{2} \sum_{i=1}^{N_e} \nabla_i^2 \right) + \left(-\frac{1}{2m_p} \sum_{i'=1}^{N_p} \nabla_{i'}^2 \right) \tag{5.5}$$

and

$$V = V_{ee} + V_{pp} + V_{ep} = \left(\sum_{i=1}^{N_e} \sum_{j>i}^{N_e} \frac{1}{\left| \mathbf{r}_i^e - \mathbf{r}_j^e \right|} \right) + \left(\sum_{i'=1}^{N_p} \sum_{j'>i'}^{N_p} \frac{1}{\left| \mathbf{r}_{i'}^p - \mathbf{r}_{j'}^p \right|} \right) + \left(-\sum_{i=1}^{N_e} \sum_{i'=1}^{N_p} \frac{1}{\left| \mathbf{r}_i^e - \mathbf{r}_{i'}^p \right|} \right). \tag{5.6}$$

Analogous to electronic DFT, we define a noninteracting reference system, in which all quantum particles (i.e., electrons and quantum protons) do not interact with each other. The ground state wavefunction of this noninteracting system is the product of electronic and nuclear Slater determinants, Φ^e and Φ^p, respectively, and is denoted as $\Psi_s\left(\mathbf{x}^e, \mathbf{x}^p\right) = \Phi^e\left(\mathbf{x}^e\right)\Phi^p\left(\mathbf{x}^p\right)$. In this expression and those that follow, $\mathbf{x}^e$ and $\mathbf{x}^p$ denote collective spatial and spin coordinates for the electrons and quantum protons, respectively. The noninteracting kinetic energy functional is given by

$$T_s\left[\rho^e, \rho^p\right] = \left\langle \Phi^e \left| T_{ee} \right| \Phi^e \right\rangle + \left\langle \Phi^p \left| T_{pp} \right| \Phi^p \right\rangle = T_s^e\left[\rho^e\right] + T_s^p\left[\rho^p\right]. \tag{5.7}$$

The Kohn-Sham scheme[26] for multicomponent DFT expresses the universal functional in terms of the noninteracting system as[19]

$$\begin{aligned} F\left[\rho^e, \rho^p\right] = T_s\left[\rho^e, \rho^p\right] + J_{ee}\left[\rho^e\right] + E_{exc}\left[\rho^e\right] + J_{pp}\left[\rho^p\right] \\ + E_{pxc}\left[\rho^p\right] + J_{ep}\left[\rho^e, \rho^p\right] + E_{epc}\left[\rho^e, \rho^p\right], \end{aligned} \tag{5.8}$$

where

$$\begin{aligned} E_{exc}\left[\rho^e\right] &= \left(T_{ee}\left[\rho^e\right] - T_s^e\left[\rho^e\right]\right) + \left(V_{ee}\left[\rho^e\right] - J_{ee}\left[\rho^e\right]\right) \\ &= T_{ec}\left[\rho^e\right] + V_{exc}\left[\rho^e\right], \end{aligned} \tag{5.9}$$

$$\begin{aligned} E_{pxc}\left[\rho^p\right] &= \left(T_{pp}\left[\rho^p\right] - T_s^p\left[\rho^p\right]\right) + \left(V_{pp}\left[\rho^p\right] - J_{pp}\left[\rho^p\right]\right) \\ &= T_{pc}\left[\rho^p\right] + V_{pxc}\left[\rho^p\right], \end{aligned} \tag{5.10}$$

and

$$\begin{aligned} E_{epc}\left[\rho^e,\rho^p\right] &= F\left[\rho^e,\rho^p\right] - T_{ee}\left[\rho^e\right] - V_{ee}\left[\rho^e\right] - T_{pp}\left[\rho^p\right] - V_{pp}\left[\rho^p\right] - J_{ep}\left[\rho^e,\rho^p\right] \\ &= T_{epc}\left[\rho^e,\rho^p\right] + V_{epc}\left[\rho^e,\rho^p\right]. \end{aligned} \tag{5.11}$$

In these expressions, the classical parts of the Coulomb interactions are given by

$$J_{ee}\left[\rho^e\right] = \frac{1}{2}\iint \mathrm{d}\mathbf{r}_1^e \mathrm{d}\mathbf{r}_2^e \, \frac{\rho^e\left(\mathbf{r}_1^e\right)\rho^e\left(\mathbf{r}_2^e\right)}{\left|\mathbf{r}_1^e - \mathbf{r}_2^e\right|} \tag{5.12}$$

and

$$J_{ep}\left[\rho^e,\rho^p\right] = -\iint \mathrm{d}\mathbf{r}_1^e \mathrm{d}\mathbf{r}_1^p \, \frac{\rho^e\left(\mathbf{r}_1^e\right)\rho^p\left(\mathbf{r}_1^p\right)}{\left|\mathbf{r}_1^e - \mathbf{r}_1^p\right|} \tag{5.13}$$

for the electron-electron and electron-proton interactions, respectively, with $J_{pp}\left[\rho^p\right]$ defined analogously for the proton-proton interaction. Furthermore, $T_{ee}\left[\rho^e\right]$ and $V_{ee}\left[\rho^e\right]$ are defined as the electronic kinetic energy and electron-electron interaction energy functionals, respectively, including electron-electron correlation effects, with analogous definitions for $T_{pp}\left[\rho^p\right]$ and $V_{pp}\left[\rho^p\right]$. In terms of the above definitions, the exact kinetic energy and potential energy functionals are expressed as

$$T\left[\rho^e,\rho^p\right] = T_s\left[\rho^e,\rho^p\right] + T_{ec}\left[\rho^e\right] + T_{pc}\left[\rho^p\right] + T_{epc}\left[\rho^e,\rho^p\right] \tag{5.14}$$

and

$$V\left[\rho^e,\rho^p\right]=J_{ee}\left[\rho^e\right]+J_{pp}\left[\rho^p\right]+J_{ep}\left[\rho^e,\rho^p\right]+V_{exc}\left[\rho^e\right]+V_{pxc}\left[\rho^p\right]+V_{epc}\left[\rho^e,\rho^p\right], \quad (5.15)$$

respectively, such that $F\left[\rho^e,\rho^p\right]=T\left[\rho^e,\rho^p\right]+V\left[\rho^e,\rho^p\right]$.

Within this formalism, the definition of the electronic exchange-correlation functional, $E_{exc}\left[\rho^e\right]$, is consistent with that used in standard electronic DFT. Thus, well-established electronic exchange-correlation functionals are expected to be suitable for use within the multicomponent DFT formalism presented here. Furthermore, the contribution from the proton exchange-correlation functional, $E_{pxc}\left[\rho^p\right]$, is assumed to be negligible due to the localized nature of protons in typical molecular systems, where only select hydrogen nuclei are treated quantum mechanically. Multiple quantum protons can be treated with a generalized Hartree-Fock approach with each quantum proton occupying a different localized spatial orbital, and the proton exchange-correlation functional can be chosen to be the diagonal proton exchange terms to eliminate the self-interaction terms.[27] Finally, the electron-proton correlation functional defined in Eq. (5.11) can be interpreted by examining Eqs. (5.14) and (5.15). The kinetic energy contribution to this functional is the kinetic energy correction arising from direct electron-proton interactions, while the potential energy contribution to this functional is the nonclassical part of the Coulombic electron-proton interactions.

Previously we developed electron-proton correlation functionals[17,20] by defining the potential energy contribution in terms of a suitable correlated electron-proton pair density, $\rho^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)$, as

$$V_{epc}\left[\rho^e,\rho^p\right]=-\iint d\mathbf{r}_1^e d\mathbf{r}_1^p \frac{\rho^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)}{\left|\mathbf{r}_1^e-\mathbf{r}_1^p\right|}-J_{ep}\left[\rho^e,\rho^p\right]. \quad (5.16)$$

Approximate electron-proton pair density expressions from two different types of explicitly

correlated wavefunctions were used to obtain two different electron-proton correlation functionals that will be discussed below. In these previous studies,[17,20] the kinetic energy contribution to the electron-proton functional was assumed to be negligible (i.e., $E_{epc}\left[\rho^e,\rho^p\right]=V_{epc}\left[\rho^e,\rho^p\right]$). In Appendix D, we derive a new electron-proton correlation functional that avoids this assumption about the kinetic energy contribution (i.e., $T_{epc}\left[\rho^e,\rho^p\right]$ is not assumed to be zero).

Given suitable exchange and correlation functionals, the Kohn-Sham prescription leads to the following coupled equations:[14]

$$\begin{aligned}\left(-\frac{1}{2}\nabla_e^2+v_{\text{eff}}^e\left(\mathbf{r}_1^e\right)\right)\psi_i^e&=\varepsilon_i^e\psi_i^e\ ,\ i=1,..,N_e/2\\ \left(-\frac{1}{2m_p}\nabla_p^2+v_{\text{eff}}^p\left(\mathbf{r}_1^p\right)\right)\psi_{i'}^p&=\varepsilon_{i'}^p\psi_{i'}^p\ ,\quad i'=1,..,N_p\end{aligned}\tag{5.17}$$

where

$$\begin{aligned}v_{\text{eff}}^e\left(\mathbf{r}_1^e\right)&=v\left(\mathbf{r}_1^e\right)-\int d\mathbf{r}_1^p\,\frac{\rho^p\left(\mathbf{r}_1^p\right)}{\left|\mathbf{r}_1^p-\mathbf{r}_1^e\right|}+\frac{\delta E_{epc}\left[\rho^e,\rho^p\right]}{\delta\rho^e\left(\mathbf{r}_1^e\right)}+\int d\mathbf{r}_2^e\,\frac{\rho^e\left(\mathbf{r}_2^e\right)}{\left|\mathbf{r}_2^e-\mathbf{r}_1^e\right|}+\frac{\delta E_{exc}\left[\rho^e\right]}{\delta\rho^e\left(\mathbf{r}_1^e\right)}\\ v_{\text{eff}}^p\left(\mathbf{r}_1^p\right)&=-v\left(\mathbf{r}_1^p\right)-\int d\mathbf{r}_1^e\,\frac{\rho^e\left(\mathbf{r}_1^e\right)}{\left|\mathbf{r}_1^e-\mathbf{r}_1^p\right|}+\frac{\delta E_{epc}\left[\rho^e,\rho^p\right]}{\delta\rho^p\left(\mathbf{r}_1^p\right)}+\int d\mathbf{r}_2^p\,\frac{\rho^p\left(\mathbf{r}_2^p\right)}{\left|\mathbf{r}_2^p-\mathbf{r}_1^p\right|}+\frac{\delta E_{pxc}\left[\rho^p\right]}{\delta\rho^p\left(\mathbf{r}_1^p\right)}.\end{aligned}\tag{5.18}$$

In these expressions, $\psi_i^e\left(\mathbf{r}_1^e\right)$ and $\psi_{i'}^p\left(\mathbf{r}_1^p\right)$ are the electron and proton Kohn-Sham spatial orbitals, respectively, that form the single-particle densities $\rho^e\left(\mathbf{r}_1^e\right)=2\sum_{i=1}^{N_e/2}\left|\psi_i^e\left(\mathbf{r}_1^e\right)\right|^2$ and $\rho^p\left(\mathbf{r}_1^p\right)=\sum_{i'=1}^{N_p}\left|\psi_{i'}^p\left(\mathbf{r}_1^p\right)\right|^2$. These Kohn-Sham equations are written for a restricted closed-shell electron system with each spatial electron orbital doubly occupied and a high-spin proton system with each spatial proton orbital singly occupied. The extension to open-shell electron systems is

straightforward. The Kohn-Sham equations can be solved iteratively to self-consistency.

B. Electron-proton correlation functionals

As mentioned above, previously we derived electron-proton correlation functionals in terms of approximate electron-proton pair density expressions obtained from explicitly correlated wavefunctions.[17,20] The EPC1 functional[17] was defined in terms of an approximate electron-proton pair density derived from the NEO-XCHF wavefunction given by[10,11]

$$\Psi_{\mathrm{XCHF}}\left(\mathbf{x}^e,\mathbf{x}^p\right)=\Phi^e\left(\mathbf{x}^e\right)\Phi^p\left(\mathbf{x}^p\right)\left[1+G\left(\mathbf{r}^e,\mathbf{r}^p\right)\right], \tag{5.19}$$

where

$$G\left(\mathbf{r}^e,\mathbf{r}^p\right)=\sum_{i=1}^{N_e}\sum_{i'=1}^{N_p}g\left(\mathbf{r}_i^e,\mathbf{r}_{i'}^p\right) \tag{5.20}$$

with Gaussian-type geminal functions defined as

$$g\left(\mathbf{r}_i^e,\mathbf{r}_{i'}^p\right)=\sum_{k=1}^{N_g}b_k e^{-\gamma_k\left|\mathbf{r}_i^e-\mathbf{r}_{i'}^p\right|^2}. \tag{5.21}$$

The electron-proton pair density for the NEO-XCHF wavefunction is[28]

$$\rho_{\mathrm{XCHF}}^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)=\frac{N_eN_p}{\left\langle\Psi_{\mathrm{XCHF}}\middle|\Psi_{\mathrm{XCHF}}\right\rangle}\left\langle\Psi_{\mathrm{XCHF}}\middle|\Psi_{\mathrm{XCHF}}\right\rangle_{-e_1p_1}. \tag{5.22}$$

In this expression and those that follow, angular brackets without subscripts denote integration over all spatial and spin coordinates, angular brackets with subscripts denote integration over only the specified spatial coordinate(s) (i.e., $\left\langle\cdots\right\rangle_{e_1}$ indicates integration over $\mathbf{r}_1^e$), and angular brackets with subscripts preceded by a minus sign denote integration over all coordinates except the specified spatial coordinate(s) (i.e., $\left\langle\cdots\right\rangle_{-e_1}$ indicates integration over all spatial and spin coordinates except $\mathbf{r}_1^e$).

In Ref. 17, the electron-proton pair density for the NEO-XCHF wavefunction was determined from Eq. (5.22). Subsequently, all terms of order g^2 and all terms that included densities other than one-particle densities were eliminated. The resulting truncated expression for the electron-proton pair density was renormalized so that $\left\langle N_e^{-1} N_p^{-1} \rho_{\mathrm{XCHF}}^{ep} \right\rangle_{e_1 p_1} = 1$. An analogous procedure was performed for the one-particle electron and proton densities, and an additional approximation described in Ref. 17 was used to derive the following approximate electron-proton pair density expression:

$$\rho_{\mathrm{XCHF}}^{ep}\left(\mathbf{r}_1^e, \mathbf{r}_1^p\right) \approx \rho^e \rho^p \left[1 + \left\langle N_e^{-1} N_p^{-1} \rho^e \rho^p g \right\rangle_{ep} - \left\langle N_e^{-1} \rho^e g \right\rangle_e - \left\langle N_p^{-1} \rho^p g \right\rangle_p + \frac{g + \left\langle N_e^{-1} \rho^e g \right\rangle_e \left\langle N_p^{-1} \rho^p g \right\rangle_p}{1 + \left\langle N_e^{-1} N_p^{-1} \rho^e \rho^p g \right\rangle_{ep}} \right]. \tag{5.23}$$

The EPC1 functional is defined by substituting this approximate electron-proton pair density into Eq. (5.16) to obtain $V_{epc}^{\mathrm{EPC1}}\left[\rho^e, \rho^p\right]$ and invoking the approximation $E_{epc}^{\mathrm{EPC1}}\left[\rho^e, \rho^p\right] = V_{epc}^{\mathrm{EPC1}}\left[\rho^e, \rho^p\right]$. Note that the derivation of the EPC1 functional requires the neglect of 24 out of 26 terms in the electron-proton pair density expression.[28]

The EPC2 functional[20] was defined in terms of an approximate electron-proton pair density derived from the alternative NEO-XCHF2 wavefunction given by[12]

$$\Psi_{\mathrm{XCHF2}}\left(\mathbf{x}^e, \mathbf{x}^p\right) = \Phi^e\left(\mathbf{x}^e\right) \Phi^p\left(\mathbf{x}^p\right) \sqrt{1 + G\left(\mathbf{r}^e, \mathbf{r}^p\right)}. \tag{5.24}$$

In Ref. 20, the electron-proton pair density for the NEO-XCHF2 wavefunction was determined from an expression analogous to that given in Eq. (5.22). The NEO-XCHF2 ansatz for the explicitly correlated wavefunction allowed us to retain all terms in the electron-proton pair density expression, provided that the two-particle electron density was approximated as

$$\rho_2^e\left(\mathbf{r}_1^e,\mathbf{r}_2^e\right)\approx\frac{1}{2}\frac{N_e-1}{N_e}\rho_1^e\left(\mathbf{r}_1^e\right)\rho_1^e\left(\mathbf{r}_2^e\right) \tag{5.25}$$

with an analogous approximation for the two-particle proton density. Following the procedure described in Ref. 20, the resulting approximate electron-proton pair density expression is

$$\begin{aligned}\rho_{\mathrm{XCHF2}}^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)\approx\rho^e\rho^p\Bigg[1+\left\langle N_e^{-1}N_p^{-1}\rho^e\rho^p g\right\rangle_{ep}-\left\langle N_e^{-1}\rho^e g\right\rangle_e-\left\langle N_p^{-1}\rho^p g\right\rangle_p \\ +\frac{g+\left\langle\rho^e g\right\rangle_e\left\langle\rho^p g\right\rangle_p}{1+\left\langle\rho^e\rho^p g\right\rangle_{ep}}\Bigg].\end{aligned} \tag{5.26}$$

The EPC2 functional is defined by substituting this approximate electron-proton pair density into Eq. (5.16) to obtain $V_{epc}^{\mathrm{EPC2}}\left[\rho^e,\rho^p\right]$ and invoking the approximation $E_{epc}^{\mathrm{EPC2}}\left[\rho^e,\rho^p\right]=V_{epc}^{\mathrm{EPC2}}\left[\rho^e,\rho^p\right]$. In contrast to the EPC1 functional, the derivation of the EPC2 functional does not require the neglect of any terms in the electron-proton pair density expression.

The derivations of both the EPC1 and EPC2 functionals are based on the assumption that $T_{epc}\left[\rho^e,\rho^p\right]=0$. In Appendix D, we demonstrate that $T_{epc}\left[\rho^e,\rho^p\right]$ can be included in the electron-proton correlation functional using the general expression

$$E_{epc}\left[\rho^e,\rho^p\right]=T_{epc}\left[\rho^e,\rho^p\right]-\iint\mathrm{d}\mathbf{r}_1^e\mathrm{d}\mathbf{r}_1^p\,\frac{\rho^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)}{\left|\mathbf{r}_1^e-\mathbf{r}_1^p\right|}-J_{ep}\left[\rho^e,\rho^p\right]. \tag{5.27}$$

In this formulation, an expression for $T_{epc}\left[\rho^e,\rho^p\right]$ is obtained using the multicomponent analog of the adiabatic connection formula.[19] Substituting the approximate NEO-XCHF2 electron-proton pair density given in Eq. (5.26) into Eq. (5.27) leads to the following expression for the EPC2-KE functional:

$$E_{epc}^{\text{EPC2-KE}}\left[\rho^e,\rho^p\right]=-\iint \mathrm{d}\mathbf{r}_1^e \mathrm{d}\mathbf{r}_1^p \frac{\rho_{\text{AC}}^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)}{\left|\mathbf{r}_1^e-\mathbf{r}_1^p\right|}-J_{ep}\left[\rho^e,\rho^p\right], \tag{5.28}$$

where

$$\begin{aligned}\rho_{\text{AC}}^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)=\rho^e\rho^p\Bigg[1+\frac{1}{2}N_e^{-1}N_p^{-1}\Omega-\frac{1}{2}\left\langle N_e^{-1}\rho^e g\right\rangle_e-\frac{1}{2}\left\langle N_p^{-1}\rho^p g\right\rangle_p \\ +\frac{\Omega-\ln\left(\Omega+1\right)}{\Omega^2}g+\frac{\Omega\left(\Omega-2\right)+2\ln\left(\Omega+1\right)}{2\Omega^3}\left\langle\rho^e g\right\rangle_e\left\langle\rho^p g\right\rangle_p\Bigg]\end{aligned} \tag{5.29}$$

with $\Omega=\left\langle\rho^e\rho^p g\right\rangle_{ep}$. The details of the derivation are presented in Appendix D. Note that the three electron-proton functionals differ only in the coefficients appearing in Eqs. (5.23), (5.26) and (5.29) for EPC1, EPC2, and EPC2-KE, respectively. Thus, the EPC2-KE functional requires no substantial computational cost beyond that of the EPC1 and EPC2 functionals.

5.3. Results and Discussion

We applied the NEO-DFT approach with the three electron-proton correlation functionals described above in conjunction with three popular electronic exchange-correlation functionals to the model system, $[\text{He-H-He}]^+$, and the isotopomers $[\text{He-D-He}]^+$ and $[\text{He-T-He}]^+$. The two helium nuclei were treated classically at a fixed distance, and the central nucleus and four electrons were treated quantum mechanically. We studied these systems with the cc-pVDZ electronic basis set,[29,30] where the electronic basis functions corresponding to the central nucleus were placed at the midpoint between the two helium nuclei. The nuclear basis set was comprised of a single *s*-type Gaussian nuclear basis function placed at the midpoint between the two helium nuclei, and the exponent was optimized variationally during the NEO-DFT calculation. We performed calculations with three different electron-proton correlation functionals, EPC1, EPC2, and EPC2-KE, in conjunction with three different electronic exchange-correlation functionals,

B3LYP, BLYP, and PBE. All calculations were performed with a modified version of the GAMESS program.[31]

To benchmark the NEO-DFT method with these various functionals, we performed three-dimensional grid calculations using the Fourier grid Hamiltonian (FGH) method[32,33] with 32 grid points in each direction. The FGH method is useful for benchmarking purposes because this model system is electronically adiabatic (i.e., the Born-Oppenheimer separation between the electrons and all nuclei is valid). In our implementation of the FGH approach, the energy is calculated for the transferring proton positioned at each point of a three-dimensional grid using standard electronic structure methods. Subsequently, the three-dimensional proton vibrational wavefunctions are determined by solving the time-independent Schrödinger equation for the proton moving on this three-dimensional potential energy surface using Fourier grid methods. Thus, in contrast to the NEO-DFT method, the FGH method assumes that the electrons respond instantaneously to the motion of the transferring proton.

Since a single *s*-type Gaussian nuclear basis function is incapable of reproducing the disparate bending and stretching vibrational frequencies of the $[He\text{-}H\text{-}He]^+$ system at its equilibrium geometry, the He–He distance was chosen to be 1.955Å, a value for which the stretching and bending frequencies were qualitatively similar at the RHF/cc-pVDZ level.[12] All calculations presented in this paper were performed at this fixed He–He distance. For benchmarking purposes, we performed FGH calculations at the RHF/cc-pVDZ level of theory and at the DFT/cc-pVDZ level of theory with the B3LYP, BLYP and PBE electronic exchange-correlation functionals.

The geminal parameters for each electron-proton correlation functional, EPC1, EPC2, and EPC2-KE, were determined following the procedure described in Ref 20. Specifically, two

Gaussian-type geminal functions were used in the expansion of Eq. (5.21), and the geminal parameters were variationally optimized using the NEO-XCHF wavefunction in Eq. (5.19) for a one-electron/one-proton model system.[12] The resulting variationally optimized geminal parameters are given in Table 2 of Ref. 12. For each electron-proton correlation functional, the geminal functions were scaled by a single parameter to ensure that NEO-DFT calculations with the electronic exchange-correlation functional chosen to be Hartree-Fock exchange reproduced the FGH vibrational stretching frequency of the $[\text{He-H-He}]^+$ system at a He–He distance of 1.955Å at the RHF/cc-pVDZ level of theory. The scaling parameters were determined to be 0.0160, 0.0224 and 0.0516 for the EPC1, EPC2, and EPC2-KE functionals, respectively. After this initial scaling of the geminal functions for each electron-proton correlation functional, the geminal parameters were fixed for all other calculations presented in this paper (i.e., for all calculations involving the various electronic exchange-correlation functionals and isotopic substitutions). Note that this scaling parameter accounts for the different treatment of electron-proton correlation in NEO-XCHF and in NEO-DFT with the three different functionals. We emphasize that the single scaling factor is present within the electron-proton functional, and the vibrational frequencies are obtained from a complete NEO-DFT calculation (i.e., the frequencies themselves are not scaled).

Table 5.1 provides the electron-electron and electron-proton correlation energies calculated for the $[\text{He-H-He}]^+$ system. The results are given for the EPC2 electron-proton functional and three different electronic exchange-correlation functionals: B3LYP, BLYP, and PBE. Here NEO-DFT(ee) indicates that only an electronic exchange-correlation functional is included (i.e., $E_{epc} = 0$), NEO-DFT(ep) denotes that an electron-proton functional is included but the electronic exchange-correlation functional is Hartree-Fock exchange, and NEO-DFT(ee,ep)

indicates that both an electronic exchange-correlation functional and an electron-proton correlation functional are included. Previously we showed that nuclear quantum effects and electronic correlation are predominantly uncoupled and that the impact of these two properties on the total energy is nearly additive.[21] In the present paper, we focus on the interplay between electron-proton and electron-electron correlation.

To examine the impact of electron-proton and electron-electron correlation on each other, we calculated each type of correlation energy with respect to the NEO-HF reference and examined their additivity with respect to the total NEO-DFT(ee,ep) energy. Specifically, we calculated the electron-electron correlation energies, $E_{\text{corr}}(\text{ee})$, defined as the difference between the NEO-DFT(ee) and NEO-HF energies, and the electron-proton correlation energy, $E_{\text{corr}}(\text{ep})$, defined as the difference between the NEO-DFT(ep) and NEO-HF energies. We also calculated the total correlation energy including electron-electron and electron-proton contributions, $E_{\text{corr}}(\text{ee,ep})$, defined as the difference between the NEO-DFT(ee,ep) and NEO-HF energies. The additivity error , which is defined as $\sigma^{\text{additivity}} = E_{\text{corr}}(\text{ee,ep}) - E_{\text{corr}}(\text{ee}) - E_{\text{corr}}(\text{ep})$, will be small when electron-electron and electron-proton correlation energies are additive contributions to the total NEO-DFT(ee,ep) energy. Table 5.1 presents the correlation energies as well as the additivity error for the three electronic exchange-correlation functionals and the EPC2 electron-proton correlation functional. These results illustrate that electron-electron and electron-proton correlation are relatively uncoupled and that the electron-electron and electron-proton correlation energies are predominantly additive effects to the total energy for these systems.

Table 5.2 presents the results of NEO-DFT calculations for the $[\text{He-H-He}]^+$ system with the electronic exchange-correlation functional chosen to be Hartree-Fock exchange. For comparison, the FGH and NEO-HF results are also provided. As observed previously,[10-12,17,20]

the NEO-HF stretching frequencies are much too high due to overlocalization of the nuclear densities. The NEO-DFT frequencies are identical to the FGH frequencies for the system with hydrogen as the central nucleus due to the scaling of the geminal parameters discussed above. Using these fixed geminal parameters, the NEO-DFT frequencies with deuterium or tritium as the central nucleus agree well with the FGH frequencies for all three electron-proton correlation functionals.

Table 5.3 demonstrates the effects of including the B3LYP, BLYP and PBE electronic exchange-correlation functionals. For each of these electronic exchange-correlation functionals, we performed NEO-DFT calculations with the three electron-proton correlation functionals, EPC1, EPC2, and EPC2-KE. The FGH stretching frequencies for the quantum nucleus reflect the change in the potential experienced by the quantum nucleus due to the inclusion of the electronic exchange-correlation functionals. Specifically, the FGH/DFT frequencies in Table 5.3 are 300-400 cm^{-1} greater than the corresponding FGH/RHF frequencies given in Table 5.2. Moreover, Table 5.3 illustrates that the NEO-DFT calculations using each of the three electron-proton correlation functionals capture this shift in vibrational stretching frequencies. For all three electronic exchange-correlation functionals, the stretching frequencies calculated with the NEO-DFT method using each of the three different electron-proton correlation functionals are in very good agreement with the corresponding FGH frequencies.

Thus, the electron-proton correlation functionals parameterized to agree with the FGH/RHF stretching frequencies also provide accurate stretching frequencies with the three different electronic exchange-correlation functionals, B3LYP, BLYP and PBE. This analysis of the stretching frequencies in Tables 5.2 and 5.3 is consistent with the analysis of the correlation energies in Table 5.1, confirming that the electron-proton and electron-electron correlation

effects are predominantly uncoupled for these systems. These results suggest that the conventional electronic exchange-correlation functionals do not need to be re-parameterized for use in this formulation of the NEO-DFT method.

The results presented in Tables 5.2 and 5.3 illustrate that the three electron-proton correlation functionals exhibit similar behavior for these model systems. As mentioned above, the derivation of the EPC1 functional requires the neglect of 24 out of 26 terms in the electron-proton pair density expression, whereas the derivation of the EPC2 functional does not require the neglect of any terms in this expression. Moreover, the EPC1 and EPC2 functionals neglect the kinetic energy contribution arising from electron-proton correlation, whereas the EPC2-KE functional includes this kinetic energy contribution. Despite these differences in the derivations, however, all three electron-proton functionals perform similarly for the model systems studied after the inclusion of the single scaling factor for the geminal functions. We expect that the results may be more disparate for larger chemical systems because the functionals scale differently with respect to the number of electrons and protons, as indicated by Eqs. (5.23), (5.26) and (5.29).

To provide further insight, we analyzed the components of the electron-proton correlation energy for the EPC2 and EPC2-KE functionals. As discussed above, these two functionals are based on the same approximate electron-proton pair density given in Eq. (5.26). The difference between these two functionals is that the EPC2 functional assumes that $T_{epc} = 0$ (i.e., $E_{epc}^{\text{EPC2}} = V_{epc}$), while the EPC2-KE functional includes T_{epc} via Eq. (D.11) (i.e., $E_{epc}^{\text{EPC2-KE}} = T_{epc} + V_{epc}$). The kinetic and potential energy components of these two electron-proton correlation functionals are given in Table 5.4 for three electronic exchange-correlation functionals. In all cases, $T_{epc} \approx -0.6 V_{epc}$ for the EPC2-KE functional. Note that the magnitude

of this prefactor would be 0.5 if the last two terms in Eq. (5.26) (i.e., the terms that are not zeroth- or first-order in g) were neglected. In addition, the values in Table 5.4 illustrate that E_{epc}^{EPC2} is virtually identical to $E_{epc}^{\mathrm{EPC2\text{-}KE}}$. The potential energy component, V_{epc}, is ~2.7 times larger for the EPC2-KE functional, but addition of $T_{epc} \approx -0.6V_{epc}$ leads to similar total electron-proton correlation energies. As mentioned above, the geminal functions were scaled for each electron-proton functional to reproduce the stretching frequency for a specific model system at a certain level of theory. The scaling factor for the geminal functions is ~2.3 times greater for the EPC2-KE functional than for the EPC2 functional, resulting in the larger magnitude of V_{epc} for the EPC-KE functional. This analysis indicates that these two functionals behave similarly due to cancellation effects between the kinetic and potential energy components of the EPC2-KE functional, in conjunction with different geminal parameters for each functional.

5.4. Conclusions

In this study, we examined the performance of the NEO-DFT method using well-established electronic exchange-correlation functionals in conjunction with three recently developed electron-proton correlation functionals. These effects were analyzed by calculating the electron-proton and electron-electron correlation energies, as well as the hydrogen vibrational frequencies, for the $[\mathrm{He\text{-}H\text{-}He}]^+$ system and the isotopomers with D and T as the central nucleus. The electron-proton and electron-electron correlation energies were found to be uncoupled and predominantly additive effects to the total energy. Grid-based calculations indicated that the hydrogen vibrational stretching frequencies increase by 300-400 cm^{-1} using the B3LYP, BLYP and PBE electronic exchange-correlation functionals compared to Hartree-Fock exchange. NEO-DFT calculations with the EPC1, EPC2, and EPC2-KE electron-proton correlation functionals

captured this shift in the vibrational stretching frequencies for all three of these electronic exchange-correlation functionals. These calculations suggest that the treatment of electron-proton correlation is predominantly uncoupled from the treatment of electron-electron correlation within the framework of the NEO-DFT approach. Future work will be directed toward testing this principle for other systems. The main consequence of predominantly additive correlation effects is that electron-proton correlation functionals and electronic exchange-correlation functionals can be developed independently and subsequently combined together without further re-parameterization.

The NEO-DFT approach is a tractable method for including nuclear quantum effects in electronic structure calculations without invoking the Born-Oppenheimer separation between the electrons and quantum nuclei. An advantage of NEO-DFT is that it treats electron-electron and electron-proton correlation effects consistently. This approach will be particularly important for studying systems with substantial electron-proton nonadiabatic effects, such as proton-coupled electron transfer reactions.[4,5] Future work will focus on developing electron-proton correlation functionals that are applicable to larger molecular systems such as inorganic catalysts and enzymes.

	$E_{\text{corr}}(\text{ee})$	$E_{\text{corr}}(\text{ep})$	$E_{\text{corr}}(\text{ee,ep})$	$\sigma^{\text{additivity}}$
B3LYP	-0.102052	-0.015067	-0.116422	0.000697
BLYP	-0.100958	-0.015067	-0.115183	0.000842
PBE	-0.076538	-0.015067	-0.090726	0.000879

Table 5.1: Electron-electron and electron-proton correlation energies for the $[\text{He-H-He}]^+$ system at a fixed He–He distance of 1.955 Å. The cc-pVDZ electronic basis set and a single *s*-type Gaussian nuclear basis function with a variationally optimized exponent were used. The NEO-DFT calculations were performed using the EPC2 electron-proton correlation functional with three different electronic exchange-correlation functionals: B3LYP, BLYP and PBE. The correlation energies $E_{\text{corr}}(\text{ee})$, $E_{\text{corr}}(\text{ep})$, and $E_{\text{corr}}(\text{ee,ep})$ are defined as the differences between the NEO-DFT(ee), NEO-DFT(ep), and NEO-DFT(ee,ep) energies, respectively, and the NEO-HF energy. The additivity error is defined as $\sigma^{\text{additivity}} = E_{\text{corr}}(\text{ee,ep}) - E_{\text{corr}}(\text{ee}) - E_{\text{corr}}(\text{ep})$. Energies are reported in atomic units.

Isotope	NEO-HF	NEO-DFT			FGH
		EPC1	EPC2	EPC2-KE	
H	3098	1191	1191	1191	1191
D	2284	805	820	824	801
T	1903	640	660	666	633

Table 5.2: Vibrational frequencies in cm^{-1} corresponding to the hydrogen vibrational stretching motion calculated with the NEO-HF, NEO-DFT, and FGH methods for the $[He\text{-}X\text{-}He]^+$ system with X = H, D, or T at a fixed He–He distance of 1.955 Å. The cc-pVDZ electronic basis set was used for all calculations, and the NEO-HF and NEO-DFT calculations utilized a single *s*-type Gaussian nuclear basis function with a variationally optimized exponent. The NEO-DFT calculations were performed using the EPC1, EPC2, and EPC2-KE electron-proton correlation functionals with the electronic exchange-correlation functional chosen to be the Hartree-Fock exchange. The FGH frequencies were obtained from the splitting between the relevant vibrational states, and the NEO frequencies were obtained from the variationally optimized exponent of the Gaussian nuclear basis function.

		EPC1	EPC2	EPC2-KE	FGH
B3LYP	H	1503	1524	1529	1519
	D	1053	1085	1094	1060
	T	856	890	890	857
BLYP	H	1589	1614	1621	1599
	D	1120	1155	1165	1124
	T	913	951	961	912
PBE	H	1634	1663	1671	1612
	D	1153	1192	1202	1132
	T	941	981	992	919

Table 5.3: Vibrational frequencies in cm^{-1} corresponding to the hydrogen vibrational stretching motion calculated with the NEO-DFT and FGH methods for the $[He\text{-}X\text{-}He]^+$ system with X = H, D, or T at a fixed He–He distance of 1.955 Å. The cc-pVDZ electronic basis set was used for all calculations, and the NEO-DFT calculations utilized a single *s*-type Gaussian nuclear basis function with a variationally optimized exponent. The NEO-DFT calculations were performed using the EPC1, EPC2, and EPC2-KE electron-proton correlation functionals with the electronic exchange-correlation functionals B3LYP, BLYP and PBE. The FGH frequencies were obtained from the splitting between the relevant vibrational states, and the NEO frequencies were obtained from the variationally optimized exponent of the Gaussian nuclear basis function.

	EPC2	EPC2-KE		
	$E_{epc} = V_{epc}$	T_{epc}	V_{epc}	$E_{epc} = T_{epc} + V_{epc}$
HF	-2.389506	3.862364	-6.251266	-2.388902
B3LYP	-2.423542	4.101573	-6.524555	-2.422982
BLYP	-2.432829	4.165889	-6.598163	-2.432274
PBE	-2.439436	4.209892	-6.648773	-2.438881

Table 5.4: Decomposition of the electron-proton correlation energy, E_{epc}, into its kinetic and potential energy components, T_{epc} and V_{epc}, respectively, for the EPC2 and EPC2-KE functionals. These values were calculated for the $[He\text{-}H\text{-}He]^+$ system at a fixed He–He distance of 1.955 Å with the cc-pVDZ electronic basis set and a single *s*-type Gaussian nuclear basis function with a variationally optimized exponent. The results are given for Hartree-Fock exchange and the B3LYP, BLYP, and PBE electronic exchange-correlation functionals. Energies are reported in atomic units.

References

(1) Y. Cha, C. J. Murray, and J. P. Klinman, *Science* **243**, 1325 (1989).
(2) M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello, *Science* **275**, 817 (1997).
(3) S. Raugei and M. L. Klein, *Journal of the American Chemical Society* **125**, 8992 (2003).
(4) S. Hammes-Schiffer and A. V. Soudackov, *Journal of Physical Chemistry B* **112**, 14108 (2008).
(5) S. Hammes-Schiffer, *Accounts of Chemical Research* **42**, 1881 (2009).
(6) S. P. Webb, T. Iordanov, and S. Hammes-Schiffer, *Journal of Chemical Physics* **117**, 4106 (2002).
(7) M. V. Pak and S. Hammes-Schiffer, *Physical Review Letters* **92**, 103002 (2004).
(8) A. Reyes, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Physics* **123**, 064104 (2005).
(9) C. Swalina and S. Hammes-Schiffer, *Journal of Physical Chemistry A* **109**, 10410 (2005).
(10) C. Swalina, M. V. Pak, A. Chakraborty, and S. Hammes-Schiffer, *Journal of Physical Chemistry A* **110**, 9983 (2006).
(11) A. Chakraborty, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Physics* **129**, 014101 (2008).
(12) C. Ko, M. V. Pak, C. Swalina, and S. Hammes-Schiffer, *Journal of Chemical Physics* **135**, 054106 (2011).
(13) J. F. Capitani, R. F. Nalewajski, and R. G. Parr, *Journal of Chemical Physics* **76**, 568 (1982).
(14) N. Gidopoulos, *Physical Review B* **57**, 2146 (1998).
(15) T. Kreibich and E. K. U. Gross, *Physical Review Letters* **86**, 2984 (2001).
(16) T. Udagawa and M. Tachikawa, *Journal of Chemical Physics* **125**, 244105 (2006).
(17) A. Chakraborty, M. V. Pak, and S. Hammes-Schiffer, *Physical Review Letters* **101**, 153001 (2008).
(18) T. Kreibich, R. van Leeuwen, and E. K. U. Gross, *Physical Review A* **78**, 022501 (2008).
(19) A. Chakraborty, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Physics* **131**, 124115 (2009).
(20) A. Sirjoosingh, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Theory and Computation* **7**, 2689 (2011).
(21) M. V. Pak, A. Chakraborty, and S. Hammes-Schiffer, *Journal of Physical Chemistry A* **111**, 4522 (2007).
(22) C. T. Lee, W. T. Yang, and R. G. Parr, *Physical Review B* **37**, 785 (1988).
(23) A. D. Becke, *Journal of Chemical Physics* **98**, 5648 (1993).
(24) J. P. Perdew, K. Burke, and M. Ernzerhof, *Physical Review Letters* **77**, 3865 (1996).
(25) M. Levy, *Proceedings of the National Academy of Sciences of the United States of America* **76**, 6062 (1979).
(26) W. Kohn and L. J. Sham, *Physical Review* **140**, 1133 (1965).
(27) B. Auer and S. Hammes-Schiffer, *Journal of Chemical Physics* **132**, 084110 (2010).
(28) A. Chakraborty and S. Hammes-Schiffer, *Journal of Chemical Physics* **129**, 204101 (2008).
(29) T. H. Dunning, *Journal of Chemical Physics* **90**, 1007 (1989).
(30) D. E. Woon and T. H. Dunning, *Journal of Chemical Physics* **100**, 2975 (1994).

(31) M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, *Journal of Computational Chemistry* **14**, 1347 (1993).
(32) C. C. Marston and G. G. Balint-Kurti, *Journal of Chemical Physics* **91**, 3571 (1989).
(33) S. P. Webb and S. Hammes-Schiffer, *Journal of Chemical Physics* **113**, 5214 (2000).

Chapter 6

Reduced Explicitly Correlated Hartree-Fock Approach within the Nuclear-Electronic Orbital Framework: Theoretical Formulation[‡]

6.1. Introduction

The inclusion of nuclear quantum effects in electronic structure calculations is important for the study of a variety of chemical systems, particularly those involving hydrogen transfer or significant hydrogen-bonding interactions.[1-3] Moreover, nonadiabatic effects between electrons and certain nuclei can be significant, as is often the case with proton-coupled electron transfer reactions.[4,5] To account for these effects, several methods that do not invoke the Born-Oppenheimer separation between electrons and specified nuclei have been developed.[6-27] One such approach is the nuclear-electronic orbital (NEO) method, which treats electrons and select nuclei quantum mechanically with an orbital-based formalism.[20-28] In this approach, typically all electrons and one or a few protons are treated quantum mechanically.

Initial methods developed within the NEO framework were based on a Hartree-Fock (HF) reference, and electron-proton correlation was incorporated using analogs of the traditional methods for describing electron-electron correlation in electronic structure theory.[27] The nuclear densities were found to be severely overlocalized due to the lack of sufficient electron-proton

[‡]Reproduced with permission from:
A. Sirjoosingh, M. V. Pak, C. Swalina, and S. Hammes-Schiffer, "Reduced explicitly correlated Hartree-Fock approach within the nuclear-electronic orbital framework: Theoretical formulation," *J. Chem. Phys.* **139**, 034102 (2013). © 2013 American Institute of Physics

correlation,[25,29-31] which is particularly significant because of the attractive electron-proton Coulomb interaction. Extensions of the NEO-HF method, such as second-order perturbation theory (NEO-MP2),[26] were determined to be insufficient for capturing enough dynamical correlation. In principle, a full configuration interaction (FCI) wavefunction built from the NEO-HF reference with a complete basis set would be exact but is clearly intractable even for small chemical systems.

More recently, explicitly correlated methods within the NEO framework have been developed.[20,22,25] This family of methods, termed NEO-XCHF (explicitly correlated Hartree-Fock) introduces correlation between electrons and quantum nuclei through Gaussian-type geminal (GTG) functions. Two explicitly correlated wavefunction-based methods (NEO-XCHF[20] and NEO-XCHF2[22]) were developed and were shown to perform reasonably well for certain model systems. However, these methods are limited in terms of computational tractability and robustness of approximations. In particular, NEO-XCHF requires the evaluation of up to five-particle integrals for a single quantum nucleus, and while NEO-XCHF2 offers significant computational advantages, certain kinetic energy contributions must be approximated. Note that a variety of explicitly correlated methods has been developed for the description of electron-electron correlation in the electronic structure community.[32-48] As discussed previously,[20,25] the majority of these approaches are not directly applicable to electron-proton correlation because of the fundamental difference between electron-electron and electron-proton correlation arising from the attractive electron-proton interaction. Specifically, the mean field NEO-HF wavefunction is not a suitable reference for perturbative treatments because of the severe overlocalization of the nuclear densities. Thus, explicit electron-proton correlation is included during the self-consistent-field procedure for optimizing the orbitals rather than

perturbatively in the NEO-XCHF and NEO-XCHF2 methods.[20,22]

An alternative approach has been the development of a multicomponent density functional theory (DFT) formalism within the NEO framework.[21,23,28] This NEO-DFT method required the development of electron-proton correlation functionals based on the pair density expressions arising from explicitly correlated NEO-XCHF wavefunctions. In this approach, standard electronic exchange-correlation functionals could be used to describe electron-electron correlation in conjunction with the more recently designed electron-proton correlation functionals. Unfortunately, the evaluation of the electron-proton correlation energy in NEO-DFT calculations is also computationally intractable for larger chemical systems.

In this paper, we develop the reduced explicitly correlated Hartree-Fock method, denoted NEO-RXCHF, within the NEO framework. Similar to NEO-XCHF, NEO-RXCHF is a wavefunction-based approach in which explicit correlation is introduced using GTG functions. The key difference between these two methods is that in RXCHF, only select electronic orbitals are explicitly correlated to the nuclear orbitals, while in XCHF, all electronic orbitals are explicitly correlated to the nuclear orbitals. The physical assumption in RXCHF is that only certain electronic orbitals (e.g., valence orbitals) need to be explicitly coupled to the quantum nuclei, while others (e.g., core orbitals) need not be correlated. Moreover, enforcing correlation of all electrons to the quantum nuclei in the same manner, as in XCHF, may be more problematic than neglecting the portions of this correlation that are expected to be insignificant, as in RXCHF.

We consider three different methods within the NEO-RXCHF formalism. The first method is based on an ansatz in which the nuclear-electronic wavefunction includes explicit coupling of only select electronic orbitals to the nuclear orbitals but is fully antisymmetric with

respect to exchange of electronic coordinates. In this case, the antisymmetrization of the wavefunction results in a computational expense similar to that of XCHF. In the second method, the electronic orbitals that are coupled to the nuclear orbitals are distinguished from the other electronic orbitals using a Hartree product formalism between the two types of electronic orbitals. In this case, the wavefunction is no longer antisymmetric with respect to exchange of all electronic coordinates, and certain electronic exchange terms are neglected, but this distinction results in a significant improvement in tractability. The third method is an extension of the second method, including approximate exchange terms between the geminal-coupled electrons and the other electrons. Each of the first two methods can be derived rigorously from an ansatz for the nuclear-electronic wavefunction, whereas the third method cannot be derived rigorously. The computational expense of the second and third methods is similar and much lower than that of the first method. For the remainder of the paper, these three methods will be denoted RXCHF-fe (full exchange), RXCHF-ne (neglect exchange), and RXCHF-ae (approximate exchange).

In the remainder of the paper, we describe the NEO-RXCHF procedure for one quantum nucleus and one geminal-coupled electronic orbital. This approach is particularly applicable to positronic systems, where all electrons and one positron (rather than one nucleus) are treated quantum mechanically, and one electronic orbital is explicitly correlated to the positronic orbital. The application of NEO-RXCHF to several positron-containing systems is reported in the accompanying paper.[49] For systems in which all electrons and a hydrogen nucleus are treated quantum mechanically, however, we expect that at least two electronic orbitals must be explicitly correlated to the quantum nuclear orbital. The extension to additional geminal-coupled electronic orbitals is straightforward and is currently under development. In this paper, we retain

the terminology for a quantum nucleus, although the formalism is also applicable to a quantum positron. For each of the first two methods, we introduce the relevant ansatz for the nuclear-electronic wavefunction and derive the corresponding energy expressions. Then we formulate the modified Fock operators required for the analog of the Hartree-Fock-Roothaan procedure for all three methods and discuss practical considerations.

6.2. Theory

We consider a system of N electrons, one quantum nucleus, which is assumed to be a proton for notational simplicity, and N_c fixed classical nuclei. The Hamiltonian in atomic units for this system is given by

$$\begin{aligned} H = &-\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \frac{1}{2m_p}\nabla_p^2 - \sum_{i=1}^{N}\sum_{A=1}^{N_c}\frac{Z_A}{\left|\mathbf{r}_i^e - \mathbf{r}_A^c\right|} + \sum_{A=1}^{N_c}\frac{Z_A}{\left|\mathbf{r}^p - \mathbf{r}_A^c\right|} \\ &+ \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{\left|\mathbf{r}_i^e - \mathbf{r}_j^e\right|} - \sum_{i=1}^{N}\frac{1}{\left|\mathbf{r}_i^e - \mathbf{r}^p\right|}, \end{aligned} \tag{6.1}$$

where $\mathbf{r}^e$, $\mathbf{r}^p$, and $\mathbf{r}^c$ denote the collective spatial coordinates of the electrons, quantum proton, and classical nuclei, respectively, m_p is the mass of the proton, and Z_A is the charge of the A^{th} classical nucleus. Note that the term corresponding to repulsion between classical nuclei is omitted from Eq. (6.1) for simplicity.

In the NEO-XCHF approach,[20,25] the nuclear-electronic wavefunction is assumed to be of the form

$$\Psi^{\text{XCHF}}\left(\mathbf{x}^e, \mathbf{x}^p\right) = \Phi^e\left(\mathbf{x}^e\right)\chi^p\left(\mathbf{x}^p\right)\left[1 + G\left(\mathbf{r}^e, \mathbf{r}^p\right)\right], \tag{6.2}$$

where $\Phi^e\left(\mathbf{x}^e\right) = \left|\chi_1^e\left(\mathbf{x}_1^e\right)\cdots\chi_N^e\left(\mathbf{x}_N^e\right)\right\rangle$ is a Slater determinant of N electronic spin orbitals, χ^p is a nuclear spin orbital, and

$$G\left(\mathbf{r}^e,\mathbf{r}^p\right)=\sum_{i=1}^{N} g\left(\mathbf{r}_i^e,\mathbf{r}^p\right) \tag{6.3}$$

with GTG functions defined as

$$g\left(\mathbf{r}_i^e,\mathbf{r}^p\right)=\sum_{k=1}^{N_{\text{gem}}} b_k e^{-\gamma_k\left|\mathbf{r}_i^e-\mathbf{r}^p\right|^2}. \tag{6.4}$$

In these expressions and those that follow, $\mathbf{x}^e$ and $\mathbf{x}^p$ denote the collective spin coordinates for the electrons and the quantum proton, respectively. The NEO-XCHF energies and modified Hartree-Fock equations have been derived and presented elsewhere.[20] Recently, this ansatz has been used to describe electron-electron correlation.[48]

In the NEO-RXCHF approach, the nuclear-electronic wavefunction is assumed to be of the form

$$\Psi^{\text{RXCHF-fe}}\left(\mathbf{x}^e,\mathbf{x}^p\right)=\frac{\chi^p\left(\mathbf{x}^p\right)}{\sqrt{N!}}\begin{vmatrix} \chi_1^e\left(\mathbf{x}_1^e\right)g\left(\mathbf{r}_1^e,\mathbf{r}^p\right) & \chi_2^e\left(\mathbf{x}_1^e\right) & \cdots & \chi_N^e\left(\mathbf{x}_1^e\right) \\ \chi_1^e\left(\mathbf{x}_2^e\right)g\left(\mathbf{r}_2^e,\mathbf{r}^p\right) & \chi_2^e\left(\mathbf{x}_2^e\right) & \cdots & \chi_N^e\left(\mathbf{x}_2^e\right) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1^e\left(\mathbf{x}_N^e\right)g\left(\mathbf{r}_N^e,\mathbf{r}^p\right) & \chi_2^e\left(\mathbf{x}_N^e\right) & \cdots & \chi_N^e\left(\mathbf{x}_N^e\right) \end{vmatrix}. \tag{6.5}$$

This ansatz contains only one electronic spin orbital that is geminal-coupled to the nuclear spin orbital, and the total wavefunction is antisymmetric with respect to the exchange of all electronic coordinates. We denote this ansatz as RXCHF-fe to emphasize that this approach includes full exchange. Note that the XCHF and RXCHF-fe wavefunctions are different in that the geminal factor is of the form $1+G$ in XCHF and G in RXCHF-fe. The reason for this choice is discussed extensively in the accompanying paper on positronic systems.[49] We emphasize that it is straightforward to interchange between the two forms of the geminal factor by choosing appropriate parameters for the GTG functions.

An approximation to the ansatz in Eq. (6.5) is given by

$$\Psi^{\text{RXCHF-ne}}\left(\mathbf{x}^e,\mathbf{x}^p\right)=\frac{\chi^p\left(\mathbf{x}^p\right)\chi_1^e\left(\mathbf{x}_1^e\right)g\left(\mathbf{r}_1^e,\mathbf{r}^p\right)}{\sqrt{(N-1)!}}\begin{vmatrix}\chi_2^e\left(\mathbf{x}_2^e\right) & \cdots & \chi_N^e\left(\mathbf{x}_2^e\right)\\ \vdots & \ddots & \vdots\\ \chi_2^e\left(\mathbf{x}_N^e\right) & \cdots & \chi_N^e\left(\mathbf{x}_N^e\right)\end{vmatrix}. \tag{6.6}$$

This ansatz also contains only one electronic spin orbital that is geminal-coupled to the nuclear spin orbital. However, it differs from the fully antisymmetric ansatz of Eq. (6.5) in that one electron (with coordinate $\mathbf{x}_1^e$) is distinguished from the others. We call this electron the "special" electron and the other electrons the "regular" electrons. We denote this ansatz as RXCHF-ne to emphasize that this approach neglects exchange between the special and regular electrons.

The energies associated with these nuclear-electronic wavefunctions are defined as expectation values over the Hamiltonian operator given in Eq. (6.1). Specifically,

$$E=\frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle}, \tag{6.7}$$

where Ψ can be obtained from Eq. (6.2), (6.5) or (6.6) to generate the XCHF, RXCHF-fe, or RXCHF-ne energy, respectively. Variation of Eq. (6.7) with respect to the spin orbitals leads to modified Hartree-Fock (HF) equations of the form

$$f_i^e\chi_i^e=\varepsilon_i^e\chi_i^e, \quad i\in\{1,\ldots N\} \tag{6.8}$$

$$f^p\chi^p=\varepsilon^p\chi^p. \tag{6.9}$$

In the standard HF method, as well as the NEO-XCHF method, the electronic Fock operators are all the same. In contrast, for the RXCHF methods, variation of the energy with respect to the geminal-coupled electronic orbital, χ_1^e, is not equivalent to variation with respect to one of the other electronic orbitals. Thus, the RXCHF methods require the solution of three coupled equations:

(1) the electronic HF equation in Eq. (6.8) with $i=1$, where we denote $f_1^e \equiv f^{e_1}$,

(2) the electronic HF equation in Eq. (6.8) with $i \neq 1$, where we denote $f_2^e = \cdots = f_N^e \equiv f^e$, and

(3) the nuclear HF equation in Eq. (6.9).

In the following subsections, we describe the evaluation of the energy expressions in Eq. (6.7) for the RXCHF-fe and RXCHF-ne methods, as well as the intermediate RXCHF-ae method. We also present an unrestricted Hartree-Fock (UHF) formalism and report working expressions for the RXCHF methods.

A. Overlap

The overlap of the RXCHF-fe wavefunction is evaluated as

$$\begin{aligned} S^{\text{RXCHF-fe}} &\equiv \left\langle \Psi^{\text{RXCHF-fe}} \middle| \Psi^{\text{RXCHF-fe}} \right\rangle \\ &= \left\langle \chi^p(p)\,\chi_1^e(1) \middle| g(1,p)^2 \middle| \chi^p(p)\,\chi_1^e(1) \right\rangle \\ &\quad - \sum_{a=2}^{N} \left\langle \chi^p(p)\,\chi_1^e(1)\,\chi_a^e(2) \middle| g(1,p)\,g(2,p) \middle| \chi^p(p)\,\chi_a^e(1)\,\chi_1^e(2) \right\rangle \\ &\equiv \left\langle \chi^p \chi_1^e \middle| g^2 \middle| \chi^p \chi_1^e \right\rangle - \sum_{a=2}^{N} \left\langle \chi^p \chi_1^e \chi_a^e \middle| g(1,p)\,g(2,p) \middle| \chi^p \chi_a^e \chi_1^e \right\rangle, \end{aligned} \tag{6.10}$$

and the overlap of the RXCHF-ne wavefunction is evaluated as

$$S^{\text{RXCHF-ne}} \equiv \left\langle \Psi^{\text{RXCHF-ne}} \middle| \Psi^{\text{RXCHF-ne}} \right\rangle = \left\langle \chi^p \chi_1^e \middle| g^2 \middle| \chi^p \chi_1^e \right\rangle. \tag{6.11}$$

In these expressions and those that follow, angular brackets denote integration over all coordinates, spin-coordinate dependence of the orbitals is abbreviated as p for $\mathbf{x}^p$ and i for $\mathbf{x}_i^e$ (and likewise for spatial coordinates appearing in the geminal functions), $g \equiv g(1,p)$, and the coordinate dependence of the spin orbitals appearing in bras or kets follows the order $p,1,2,\ldots$,

e.g., $\left|\chi^p \chi_1^e \chi_a^e\right\rangle \equiv \left|\chi^p(p)\chi_1^e(1)\chi_a^e(2)\right\rangle$.

B. Energy

The energy associated with the RXCHF-fe wavefunction can be expressed as

$$E^{\text{RXCHF-fe}} = \frac{E_1 + E_2 + E_3 + E_4}{S^{\text{RXCHF-fe}}}, \tag{6.12}$$

where the overlap is given in Eq. (6.10). We omit the details of the derivation of the individual energy components here but include an example derivation of the one-electron RXCHF-fe energy in Section E.5.

The first-order terms are given by

$$E_1 = \left\langle \chi^p \chi_1^e \left|\Omega_1(p,1)\right| \chi^p \chi_1^e \right\rangle, \tag{6.13}$$

where

$$\Omega_1(p,1) = g(1,p)\left[h^e(1) + h^p(p) + V_{ep}(p,1)\right] g(1,p). \tag{6.14}$$

In this expression and those that follow, $h^e(i)$ is the core electronic operator for the i^{th} electron, $h^p(p)$ is the core nuclear operator, and $V_{ep}(p,i) = -\left|\mathbf{r}^p - \mathbf{r}_i^e\right|^{-1}$ is the Coulomb attraction operator between the quantum nucleus and the i^{th} electron.

The second-order terms are given by

$$E_2 = \sum_{a=2}^{N}\left(\left\langle \chi^p \chi_1^e \chi_a^e \left|\Omega_2^{(1)}(p,1,2)\right| \chi^p \chi_1^e \chi_a^e \right\rangle - \left\langle \chi^p \chi_1^e \chi_a^e \left|\Omega_2^{(2)}(p,1,2)\right| \chi^p \chi_a^e \chi_1^e \right\rangle\right), \tag{6.15}$$

where

$$\Omega_2^{(1)}(p,1,2) = g(1,p)\left[h^e(2) + V_{ee}(1,2) + V_{ep}(p,2)\right] g(1,p), \tag{6.16}$$

$$\Omega_2^{(2)}(p,1,2) = g(1,p)\left[h^e(1) + h^e(2) + h^p(p) + V_{ee}(1,2) + V_{ep}(p,1) + V_{ep}(p,2)\right] g(2,p). \tag{6.17}$$

In this expression and those that follow, $V_{ee}(i,j)=\left|\mathbf{r}_i^e-\mathbf{r}_j^e\right|^{-1}$ is the Coulomb repulsion operator between the i^{th} and j^{th} electrons.

The third-order terms are given by

$$\begin{aligned}E_3=&\sum_{a=2}^{N}\sum_{b=2}^{N}\\&\left(\frac{1}{2}\left\langle\chi^p\chi_1^e\chi_a^e\chi_b^e\left|\Omega_3^{(1)}(p,1,2,3)\right|\chi^p\chi_1^e\chi_a^e\chi_b^e\right\rangle-\frac{1}{2}\left\langle\chi^p\chi_1^e\chi_a^e\chi_b^e\left|\Omega_3^{(1)}(p,1,2,3)\right|\chi^p\chi_1^e\chi_b^e\chi_a^e\right\rangle\right.\\&\left.-\left\langle\chi^p\chi_1^e\chi_a^e\chi_b^e\left|\Omega_3^{(2)}(p,1,2,3)\right|\chi^p\chi_a^e\chi_1^e\chi_b^e\right\rangle+\left\langle\chi^p\chi_1^e\chi_a^e\chi_b^e\left|\Omega_3^{(2)}(p,1,2,3)\right|\chi^p\chi_b^e\chi_1^e\chi_a^e\right\rangle\right),\end{aligned}\tag{6.18}$$

where

$$\Omega_3^{(1)}(p,1,2,3)=g(1,p)\left[V_{ee}(2,3)\right]g(1,p),\tag{6.19}$$

$$\Omega_3^{(2)}(p,1,2,3)=g(1,p)\left[h^e(3)+V_{ee}(1,3)+V_{ee}(2,3)+V_{ep}(p,3)\right]g(2,p).\tag{6.20}$$

The fourth-order terms are given by

$$\begin{aligned}E_4=&\sum_{a=2}^{N}\sum_{b\neq a}^{N}\sum_{c\neq a}^{N}\left(-\frac{1}{2}\left\langle\chi^p\chi_1^e\chi_a^e\chi_b^e\chi_c^e\left|\Omega_4(p,1,2,3,4)\right|\chi^p\chi_a^e\chi_1^e\chi_b^e\chi_c^e\right\rangle\right.\\&+\frac{1}{2}\left\langle\chi^p\chi_1^e\chi_a^e\chi_b^e\chi_c^e\left|\Omega_4(p,1,2,3,4)\right|\chi^p\chi_a^e\chi_1^e\chi_c^e\chi_b^e\right\rangle\\&+\left\langle\chi^p\chi_1^e\chi_a^e\chi_b^e\chi_c^e\left|\Omega_4(p,1,2,3,4)\right|\chi^p\chi_b^e\chi_1^e\chi_a^e\chi_c^e\right\rangle\\&\left.-\left\langle\chi^p\chi_1^e\chi_a^e\chi_b^e\chi_c^e\left|\Omega_4(p,1,2,3,4)\right|\chi^p\chi_b^e\chi_1^e\chi_c^e\chi_a^e\right\rangle\right),\end{aligned}\tag{6.21}$$

where

$$\Omega_4(p,1,2,3,4)=g(1,p)\left[V_{ee}(3,4)\right]g(2,p).\tag{6.22}$$

Because Eq. (6.21) contains restricted summations, the Fock operators for all of the regular electrons would not be identical, and nonstandard density matrices would need to be introduced to solve each HF equation separately. For the remainder of this paper, we assume that in the RXCHF-fe formulation, the system of interest contains at most three electrons so that the energy

expression in Eq. (6.12) reduces to

$$E^{\text{RXCHF-fe}} = \frac{E_1 + E_2 + E_3}{S^{\text{RXCHF-fe}}}. \tag{6.23}$$

Development of the RXCHF-fe formalism for the general case of four or more electrons is left to future studies.

The energy associated with the RXCHF-ne wavefunction can be expressed as

$$E^{\text{RXCHF-ne}} = E_{\text{core}} + E_{ee} + E_{G1} + E_{G2}, \tag{6.24}$$

where

$$E_{\text{core}} = \sum_{a=2}^{N} \left\langle \chi_a^e \left| h^e(1) \right| \chi_a^e \right\rangle \tag{6.25}$$

is the usual core electronic energy for the regular electrons and

$$E_{ee} = \frac{1}{2} \sum_{a=2}^{N} \sum_{b=2}^{N} \left(\left\langle \chi_a^e \chi_b^e \left| \frac{1}{r_{12}} \right| \chi_a^e \chi_b^e \right\rangle - \left\langle \chi_a^e \chi_b^e \left| \frac{1}{r_{12}} \right| \chi_b^e \chi_a^e \right\rangle \right) \tag{6.26}$$

represents the usual electronic Coulomb and exchange terms for the regular electrons. The geminal contributions are defined as

$$E_{G1} = \frac{1}{S^{\text{RXCHF-ne}}} \left\langle \chi^p \chi_1^e \left| \Omega_1(p,1) \right| \chi^p \chi_1^e \right\rangle, \tag{6.27}$$

where Ω_1 is given by Eq. (6.14), and

$$E_{G2} = \frac{1}{S^{\text{RXCHF-ne}}} \sum_{a=2}^{N} \left\langle \chi^p \chi_1^e \chi_a^e \left| \Omega_2(p,1,2) \right| \chi^p \chi_1^e \chi_a^e \right\rangle, \tag{6.28}$$

where

$$\Omega_2(p,1,2) = g(1,p) \left[V_{ee}(1,2) + V_{ep}(p,2) \right] g(1,p). \tag{6.29}$$

In contrast to the RXCHF-fe formulation, the RXCHF-ne energy expression does not include restricted summations. Therefore, the subsequent formulation of the RXCHF-ne method is

applicable to systems with any number of electrons.

By comparing Eqs. (6.12) and (6.24), we observe that the RXCHF-ne method offers significant advantages in computational tractability over the RXCHF-fe method. Specifically, by distinguishing the one electron that is geminal-coupled to the quantum nucleus from the other electrons, the RXCHF-ne energy requires the evaluation of only up to three-particle integrals, whereas the energy associated with the fully antisymmetric ansatz of RXCHF-fe requires evaluation of up to five-particle integrals, as is the case with NEO-XCHF.

With the motivation of maintaining the RXCHF-ne level of tractability while accounting for exchange effects between special and regular electrons, we propose the RXCHF-ae method, which includes approximate exchange. In this method, the energy is expressed as

$$E^{\text{RXCHF-ae}} = E^{\text{RXCHF-ne}} + E_{\text{ex}}, \tag{6.30}$$

with the approximate exchange term given by

$$E_{\text{ex}} = -\frac{1}{S^{\text{RXCHF-ne}}} \sum_{a=2}^{N} \left\langle \chi^p \chi_1^e \chi_a^e \left| \Omega_{\text{ex}}(p,1,2) \right| \chi^p \chi_a^e \chi_1^e \right\rangle, \tag{6.31}$$

where

$$\Omega_{\text{ex}}(p,1,2) = g(1,p)\left[V_{ee}(1,2)\right] g(2,p). \tag{6.32}$$

Note that this energy contribution is a part of the second-order energy of the RXCHF-fe expression given in Eq. (6.15). In the RXCHF-ae method, the energy in Eq. (6.30) is minimized with respect to the spin orbitals. Due to the addition of E_{ex}, however, the RXCHF-ae energy does not arise from a specific wavefunction, and approximations need to be invoked to evaluate expectation values of other operators. As shown in the accompanying paper,[49] for a given set of orbital coefficients, the approximate exchange term given in Eq. (6.31) accounts for more than 99% of the difference in RXCHF-fe and RXCHF-ne energies for the positron-lithium system,

where all electrons and the positron are treated quantum mechanically.

C. Fock operators

As described above, three modified HF equations corresponding to the nuclear orbital, the geminal-coupled electronic orbital, and the regular electronic orbitals must be solved self-consistently in the variational procedure. Specifically, the three spin-coordinate-dependent equations are

$$f^{p}\left(\mathbf{x}^{p}\right)\chi^{p}\left(\mathbf{x}^{p}\right)=\varepsilon^{p}\chi^{p}\left(\mathbf{x}^{p}\right), \tag{6.33}$$

$$f^{e_1}\left(\mathbf{x}_1^{e}\right)\chi_1^{e}\left(\mathbf{x}_1^{e}\right)=\varepsilon_{11}^{e}\chi_1^{e}\left(\mathbf{x}_1^{e}\right)+\sum_{a=2}^{N}\varepsilon_{a1}\chi_a^{e}\left(\mathbf{x}_1^{e}\right), \tag{6.34}$$

$$f^{e}\left(\mathbf{x}_2^{e}\right)\chi_a^{e}\left(\mathbf{x}_2^{e}\right)=\varepsilon_a^{e}\chi_a^{e}\left(\mathbf{x}_2^{e}\right), \quad 2\le a\le N. \tag{6.35}$$

Note that Eq. (6.34) includes terms with off-diagonal Lagrange multipliers to ensure that the geminal-coupled orbital is constrained to be orthogonal to the regular electronic orbitals. These terms can rigorously be eliminated through the orthogonalization scheme described in Section 6.2F.

The Fock operators in these equations are determined for the various methods by varying the energy expressions in Eqs. (6.23), (6.24), and (6.30) with respect to the appropriate orbitals. The resulting expressions for RXCHF-fe, RXCHF-ne, and RXCHF-ae are given in Sections E.2.1, E.3.1, and E.4.1, respectively.

D. Spatial Fock operators

To obtain the spatial Fock operators and their corresponding eigenvalue equations, we integrate Eqs. (6.33) to (6.35) over the spins of the electrons as well as the spin of the quantum

nucleus. For all RXCHF methods, the spatial part of the quantum nuclear spin orbital is denoted ψ^p without specifying the spin of this particle because integration over the spin of this orbital always leads to unity in expectation value expressions for spin-independent operators. Moreover, without loss of generality, we assume that the spin of the geminal-coupled orbital, χ_1^e, is α:

$$\chi_1^e\left(\mathbf{x}_1^e\right)=\psi_1^e\left(\mathbf{r}_1^e\right)\alpha\left(\omega_1^e\right), \tag{6.36}$$

where ψ_1^e is the spatial part of the geminal-coupled spin orbital. Finally, following the unrestricted Hartree-Fock (UHF) formalism, we partition the regular electronic spin orbitals as

$$\left\{\chi_a^e\left(\mathbf{r}_a^e,\omega_a^e\right)\right\}=\left\{\psi_a^\alpha\left(\mathbf{r}_a^e\right)\alpha\left(\omega_a^e\right)\right\}\cup\left\{\psi_a^\beta\left(\mathbf{r}_a^e\right)\beta\left(\omega_a^e\right)\right\}\equiv \mathrm{A}\cup\mathrm{B}, \tag{6.37}$$

where $\left|\mathrm{A}\right|=N_\alpha$ and $\left|\mathrm{B}\right|=N_\beta$ such that $N_\alpha+N_\beta=N-1$. Note that the index *a* spans a different range in the context of the spatial orbitals: for ψ_a^α, $1\le a\le N_\alpha$, and for ψ_a^β, $1\le a\le N_\beta$.

The integration of Eqs. (6.33) and (6.34) over spin results in the corresponding spatial-coordinate analogs. The integration of Eq. (6.35) over α or β spin leads to two different sets of modified HF equations,

$$f^\alpha\left(\mathbf{r}_2^e\right)\psi_a^\alpha\left(\mathbf{r}_2^e\right)=\varepsilon_a^\alpha\psi_a^\alpha\left(\mathbf{r}_2^e\right),\ \ 1\le a\le N_\alpha \tag{6.38}$$

$$f^\beta\left(\mathbf{r}_2^e\right)\psi_a^\beta\left(\mathbf{r}_2^e\right)=\varepsilon_a^\beta\psi_a^\beta\left(\mathbf{r}_2^e\right),\ \ 1\le a\le N_\beta. \tag{6.39}$$

Thus, the UHF formalism applied to the RXCHF methods requires the solution of four coupled modified HF equations. The resulting spatial Fock operator expressions for RXCHF-fe, RXCHF-ne, and RXCHF-ae are given in Sections E.2.2, E.3.2, and E.4.2, respectively. Note that this unrestricted approach could potentially lead to spin contamination. To address this issue, we are also developing a restricted open-shell formalism analogous to ROHF methods in electronic structure theory.

E. Modified Hartree-Fock-Roothaan expressions

We now define the atomic orbital (AO) basis sets and report the analogs of Hartree-Fock-Roothaan equations for the RXCHF methods. We expand the nuclear spatial orbital in a set of N_{pbf} nuclear basis functions as

$$\psi^{p}\left(\mathbf{r}^{p}\right)=\sum_{\mu'=1}^{N_{\text{pbf}}}C_{\mu'}^{p}\phi_{\mu'}^{p}\left(\mathbf{r}^{p}\right). \tag{6.40}$$

In the most general case, we expand all of the electronic spatial orbitals in a common set of N_{ebf} electronic basis functions. The geminal-coupled electronic orbital is expanded as

$$\psi_{1}^{e}\left(\mathbf{r}_{1}^{e}\right)=\sum_{\mu''=1}^{N_{\text{ebf}}}C_{\mu''}^{e_1}\phi_{\mu''}^{e}\left(\mathbf{r}_{1}^{e}\right), \tag{6.41}$$

while the other electronic orbitals are expanded as

$$\psi_{a}^{\alpha}\left(\mathbf{r}_{2}^{e}\right)=\sum_{\mu=1}^{N_{\text{ebf}}}C_{\mu a}^{\alpha}\phi_{\mu}^{e}\left(\mathbf{r}_{2}^{e}\right), \tag{6.42}$$

$$\psi_{a}^{\beta}\left(\mathbf{r}_{2}^{e}\right)=\sum_{\mu=1}^{N_{\text{ebf}}}C_{\mu a}^{\beta}\phi_{\mu}^{e}\left(\mathbf{r}_{2}^{e}\right). \tag{6.43}$$

Note that substantial computational savings could be gained by restricting the basis set of the geminal-coupled electronic orbital. Utilizing knowledge about the chemical environment of the quantum nucleus, only a subset of the AO basis functions centered on particular atoms could be included in the expansion of Eq. (6.41). For example, a hydrogen-bonding interface with the hydrogen nucleus treated quantum mechanically may be well-described by geminal-coupled electronic orbitals comprised of AOs on only the donor, acceptor, and hydrogen atoms at the interface.

The spatial Fock operators discussed in the previous subsection can be defined in terms

of these quantities to provide working expressions of the form $\mathbf{FC} = \mathbf{SCE}$ for appropriate modified Fock matrices **F** in the AO basis. The expressions for the RXCHF-fe, RXCHF-ne, and RXCHF-ae Fock matrices are given in Sections E.2.3, E.3.3, and E.4.3, respectively.

F. Orthogonalization Scheme

In the RXCHF schemes described above, a modified HF equation with Fock operator f^{e_1} for the geminal-coupled electronic orbital is solved separately from the modified HF equations for the regular alpha and beta electronic orbitals with Fock operators f^{α} and f^{β}, respectively. At each iteration of the self-consistent-field (SCF) procedure, the geminal-coupled electronic orbital, χ_1^e, which is assumed to have spin α, is not necessarily orthogonal to the regular alpha electronic orbitals, $\chi_a^e \in \mathrm{A}$. (Due to spin orthogonality, χ_1^e is always orthogonal to all beta electronic orbitals, $\chi_a^e \in \mathrm{B}$.) To ensure that ψ_1^e remains orthogonal to the spatial part of the regular alpha electronic orbitals, we adopt a modification of the orthogonality constrained basis set expansion (OCBSE) method.[50]

In the original formulation of the OCBSE method applied to open-shell restricted HF schemes, one set of orbitals is constrained to be orthogonal to another set by expanding each in an appropriate basis set spanning the orthogonal complement of the respective occupied spaces.[50,51] The coupled HF equations are projected onto these basis sets in order to eliminate off-diagonal Lagrange multipliers to allow the use of standard matrix equations during the SCF procedure. The main drawback of the OCBSE procedure is that the solutions of the coupled HF equations are calculated over a more reduced space than would be the case if each were expanded fully in identical AO basis sets.

For the RXCHF methods, we use a slightly modified version of the OCBSE procedure, where only the HF equation for the geminal-coupled electronic orbital is projected onto the orthogonal complement of the space spanned by the occupied regular electronic orbitals. This procedure is described below for iteration i assuming orthogonal orbitals from the previous iteration. We adopt a slightly different notation in this subsection for clarity.

We define $\psi^{(i-1)}$ as the geminal-coupled electronic orbital obtained from the $i-1^{\text{th}}$ iteration and $\left\{\psi_a^{(i-1)} : 1 \le a \le N_\alpha\right\}$ as the occupied regular alpha electronic orbitals obtained from the $i-1^{\text{th}}$ iteration. From the previous iteration, we have density matrices that enable the calculation of the matrices $\mathbf{F}^{e_1}$ and $\mathbf{F}^{\alpha}$, which are defined in the AO basis in Appendix E. We solve the usual Hartree-Fock-Roothaan equation for the regular alpha electronic orbitals in the AO basis,

$$\mathbf{F}^{\alpha}\mathbf{C}^{\alpha} = \mathbf{S}\mathbf{C}^{\alpha}\mathbf{E}^{\alpha}, \tag{6.44}$$

which results in the alpha electronic orbitals of the current iteration, $\left\{\psi_a^{(i)} : 1 \le a \le N^{\text{ebf}}\right\}$, where the orbitals are ordered such that the first N^{α} are occupied and the remaining $N^{\text{ebf}} - N_\alpha$ are virtual. By the diagonalization procedure used to solve Eq. (6.44), this set of occupied and virtual alpha electronic orbitals is orthonormal. Note that the Hartree-Fock-Roothaan equation for the regular electronic orbitals is free of any constraints depending on the geminal-coupled electronic orbital, i.e., no off-diagonal Lagrange multipliers appear in this equation. The energy is minimized with respect to the uncoupled electronic orbitals without enforcing orthogonality to the geminal-coupled electronic orbital. This constraint is applied only to the Hartree-Fock-Roothaan equation for the geminal-coupled electronic orbital.

The next step is to transform the corresponding Hartree-Fock-Roothaan equation for the

geminal-coupled electronic orbital in the AO basis,

$$\mathbf{F}^{e_1}\mathbf{C}^{e_1} = \mathbf{S}\mathbf{C}^{e_1}\mathbf{E}^{e_1}, \tag{6.45}$$

to an orthonormal basis (denoted $B^{(i)}$) that spans a space orthogonal to the space spanned by the occupied regular electronic orbitals $\left\{\psi_a^{(i)} : 1 \le a \le N_\alpha\right\}$. We define this basis as

$$B^{(i)} = \left\{\psi_a^{(i)} : N_\alpha < a \le N^{\text{ebf}}\right\}, \tag{6.46}$$

which is comprised of the virtual electronic orbitals for the regular electrons in the current iteration. Transforming Eq. (6.45) before diagonalization provides a solution for the geminal-coupled electronic orbital, $\psi^{(i)}$, that is orthogonal to all of the occupied regular electronic orbitals, $\left\{\psi_a^{(i)} : 1 \le a \le N_\alpha\right\}$.

Following the formalism presented in Refs. 50 and 51, the projection onto $B^{(i)}$ rigorously eliminates the off-diagonal Lagrange multipliers in the Hartree-Fock-Roothaan equation for the geminal-coupled electronic orbital. Thus, this procedure minimizes the energy with respect to variations in the geminal-coupled electronic orbital subject to the constraint that this orbital is orthogonal to all of the occupied regular electronic orbitals. Note that alternative orthogonalization schemes may be used within the framework of the RXCHF methods. An alternative scheme used in some of our calculations is provided in Section E.6.

6.3. Discussion and Practical Considerations

The evaluation of the geminal integrals in Eqs. (E.17) – (E.27) is discussed in detail in previous papers.[20] In particular, we use an extension of the McMurchie-Davidson approach[52] to evaluate three-, four- and five-particle geminal integrals.[53] More efficient integral schemes are currently being implemented and tested, including resolution of identity (RI) approximations, the

Rys quadrature approach,[54] and the new range of tensor hypercontraction methods.[55-57] These schemes are expected to offer significant advantages in computational tractability for the XCHF and RXCHF methods.

The choice of GTG function parameters for positronic systems is discussed in the accompanying paper.[49] In particular, we demonstrate that geminal parameters obtained from a single optimization for a one-electron-one-positron model system are suitable for the various other positron-containing systems studied. We expect that this transferability will be applicable to other types of systems (e.g., proton-containing systems) due to the local nature of the electron-nucleus interaction, although the GTG parameters for electron-proton interactions will be different from those for electron-positron interactions. Thus, a new set of GTG parameters must be obtained for each type of quantum particle, but these parameters should be transferable among systems containing this type of quantum particle.

The implementation of RXCHF-fe described above is limited to $N < 4$ electrons and still requires the evaluation of four-particle integrals. As discussed above, the computational expense is comparable to XCHF calculations for systems of the same size. Nevertheless, RXCHF-fe possesses fundamentally different underlying physical assumptions. In the XCHF method, all electronic orbitals are explicitly correlated to the nuclear orbitals in the same manner using the same GTG parameters. In this case, the geminal functions are often used to account for interactions other than the short-ranged electron-nucleus dynamical correlation. The RXCHF-fe method avoids this problem by explicitly correlating only the relevant electrons to the quantum nuclei, ensuring that the geminal parameters are used predominantly to describe the short-ranged electron-nucleus interaction. Thus, typically the XCHF wavefunction is more globally optimized, but the RXCHF-fe wavefunction provides a more accurate description of the short-ranged

electron-nucleus interaction.

The introduction of the RXCHF-ne and RXCHF-ae approximate methods provides substantially more efficient approaches based on the same underlying physical principles as the more rigorous RXCHF-fe method. These methods involve the calculation of only up to three-particle integrals and avoid the four- and five-particle integrals required for the XCHF and RXCHF-fe methods. Combined with more efficient integral techniques and localized basis sets for the geminal-coupled electronic orbitals, the RXCHF-ne and RXCHF-ae methods provide a promising prospect for studying larger chemical systems. Their accuracy can be tested for smaller systems by comparison to RXCHF-fe to determine the effects of approximating the full electronic exchange interactions. Developing and testing additional schemes to approximate these exchange interactions is an important direction for future research. Extensions of frozen-core and effective core potential methods are also promising directions.

The accompanying paper[49] demonstrates the application of the RXCHF methods to several positron-containing molecular species, where the electrons and positrons are treated quantum mechanically. We observe that all RXCHF methods outperform XCHF and that the RXCHF-ne and RXCHF-ae methods are excellent approximations to the RXCHF-fe method. These results provide an optimistic outlook for the application of efficient and accurate implementations of the RXCHF methods to larger chemical systems. Current research directions focus on extending the formalism presented in this paper to the case of more than one geminal-coupled electronic orbital. A major goal is to develop RXCHF methods for the investigation of proton-containing molecular species, where the electrons and protons are treated quantum mechanically.

6.4. Conclusions

In this paper, we presented the RXCHF approach as an alternative to the previously developed XCHF approach for including explicit electron-nucleus correlation within the NEO framework. The primary paradigm shift lies in restricting the explicit electron-nucleus correlation to only select electronic orbitals. The RXCHF-fe, RXCHF-ne, and RXCHF-ae methods were defined and compared in terms of the degree of electronic exchange contributions and computational expense. Working expressions for each method were derived and reported in the form of modified Fock operators for coupled sets of Hartree-Fock-Roothaan equations. The RXCHF-ne and RXCHF-ae methods offer substantial computational savings over the XCHF approach. Based on the success for positron-containing molecular species, as described in the accompanying paper,[49] the outlook for application of the RXCHF methods to larger chemical systems is promising. A particularly important direction is the application of the RXCHF methods to chemical systems in which nuclear quantum effects are important, especially those involving proton-coupled electron transfer reactions, which often exhibit non-Born-Oppenheimer effects between the electrons and transferring proton(s).[4,5] In these cases, more than a single electronic orbital will be explicitly correlated to the quantum nucleus. Extensions to enable such calculations are currently under development.

References

(1) Y. Cha, C. J. Murray, and J. P. Klinman, *Science* **243**, 1325 (1989).
(2) S. Raugei and M. L. Klein, *Journal of the American Chemical Society* **125**, 8992 (2003).
(3) M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello, *Science* **275**, 817 (1997).
(4) S. Hammes-Schiffer, *Energy & Environmental Science* **5**, 7696 (2012).
(5) A. Sirjoosingh and S. Hammes-Schiffer, *Journal of Physical Chemistry A* **115**, 2367 (2011).
(6) A. D. Bochevarov, E. F. Valeev, and C. D. Sherrill, *Molecular Physics* **102**, 111 (2004).
(7) S. Bubin, M. Pavanelo, W. C. Tung, K. L. Sharkey, and L. Adamowicz, *Chemical Reviews* **113**, 36 (2013).
(8) M. Goli and S. Shahbazian, *Theoretical Chemistry Accounts* **131**, 1365 (2013).
(9) M. Hoshino, H. Nishizawa, and H. Nakai, *Journal of Chemical Physics* **135**, 024111 (2011).
(10) E. Kamarchik and D. A. Mazziotti, *Physical Review A* **75**, 013203 (2007).
(11) E. Kamarchik and D. A. Mazziotti, *Physical Review A* **79**, 012502 (2009).
(12) T. Kreibich and E. K. U. Gross, *Physical Review Letters* **86**, 2984 (2001).
(13) T. Kreibich, R. van Leeuwen, and E. K. U. Gross, *Physical Review A* **78**, 022501 (2008).
(14) E. Matyus and M. Reiher, *Journal of Chemical Physics* **137**, 024104 (2012).
(15) H. Nakai, *International Journal of Quantum Chemistry* **86**, 511 (2002).
(16) H. Nakai, K. Sodeyama, and M. Hoshino, *Chemical Physics Letters* **345**, 118 (2001).
(17) Y. Shigeta, H. Nagao, K. Nishikawa, and K. Yamaguchi, *Journal of Chemical Physics* **111**, 6171 (1999).
(18) T. Shimazaki and M. Kubo, *Chemical Physics Letters* **525-26**, 134 (2012).
(19) M. Tachikawa, K. Mori, H. Nakai, and K. Iguchi, *Chemical Physics Letters* **290**, 437 (1998).
(20) A. Chakraborty, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Physics* **129**, 014101 (2008).
(21) A. Chakraborty, M. V. Pak, and S. Hammes-Schiffer, *Physical Review Letters* **101**, 153001 (2008).
(22) C. Ko, M. V. Pak, C. Swalina, and S. Hammes-Schiffer, *Journal of Chemical Physics* **135**, 054106 (2011).
(23) A. Sirjoosingh, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Theory and Computation* **7**, 2689 (2011).
(24) J. H. Skone, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Physics* **123**, 134108 (2005).
(25) C. Swalina, M. V. Pak, A. Chakraborty, and S. Hammes-Schiffer, *Journal of Physical Chemistry A* **110**, 9983 (2006).
(26) C. Swalina, M. V. Pak, and S. Hammes-Schiffer, *Chemical Physics Letters* **404**, 394 (2005).
(27) S. P. Webb, T. Iordanov, and S. Hammes-Schiffer, *Journal of Chemical Physics* **117**, 4106 (2002).
(28) A. Sirjoosingh, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Physics* **136**, 174114 (2012).
(29) M. V. Pak and S. Hammes-Schiffer, *Physical Review Letters* **92**, 103002 (2004).

(30) M. V. Pak, C. Swalina, S. P. Webb, and S. Hammes-Schiffer, *Chemical Physics* **304**, 227 (2004).
(31) C. Swalina and S. Hammes-Schiffer, *Journal of Physical Chemistry A* **109**, 10410 (2005).
(32) S. F. Boys, *Proceedings of the Royal Society of London Series A - Mathematical and Physical Sciences* **258**, 402 (1960).
(33) M. Cafiero, S. Bubin, and L. Adamowicz, *Physical Chemistry Chemical Physics* **5**, 1491 (2003).
(34) W. Cencek, J. Komasa, and J. Rychlewski, *Chemical Physics Letters* **246**, 417 (1995).
(35) W. Cencek and J. Rychlewski, *Journal of Chemical Physics* **98**, 1252 (1993).
(36) S. Kedzuch, M. Milko, and J. Noga, *International Journal of Quantum Chemistry* **105**, 929 (2005).
(37) W. Klopper and W. Kutzelnigg, *Journal of Chemical Physics* **94**, 2020 (1991).
(38) T. Korona, H. L. Williams, R. Bukowski, B. Jeziorski, and K. Szalewicz, *Journal of Chemical Physics* **106**, 5109 (1997).
(39) W. Kutzelnigg and W. Klopper, *Journal of Chemical Physics* **94**, 1985 (1991).
(40) F. R. Manby, H. J. Werner, T. B. Adler, and A. J. May, *Journal of Chemical Physics* **124** (2006).
(41) D. A. Mazziotti, *Journal of Chemical Physics* **112**, 10125 (2000).
(42) B. J. Persson and P. R. Taylor, *Journal of Chemical Physics* **105**, 5915 (1996).
(43) V. A. Rassolov, *Journal of Chemical Physics* **117**, 5978 (2002).
(44) S. Ten-no, *Chemical Physics Letters* **330**, 169 (2000).
(45) V. Termath, W. Klopper, and W. Kutzelnigg, *Journal of Chemical Physics* **94**, 2002 (1991).
(46) E. F. Valeev, *Chemical Physics Letters* **395**, 190 (2004).
(47) E. F. Valeev, *Journal of Chemical Physics* **125** (2006).
(48) S. A. Varganov and T. J. Martinez, *Journal of Chemical Physics* **132**, 054103 (2010).
(49) A. Sirjoosingh, M. V. Pak, C. Swalina, and S. Hammes-Schiffer, *Journal of Chemical Physics* **139**, 034103 (2013).
(50) W. J. Hunt, T. H. Dunning, Jr., and W. A. Goddard, *Chemical Physics Letters* **3**, 606 (1969).
(51) S. Krebs, *Computer Physics Communications* **116**, 137 (1999).
(52) L. E. McMurchie and E. R. Davidson, *Journal of Computational Physics* **26**, 218 (1978).
(53) B. J. Persson and P. R. Taylor, *Theoretical Chemistry Accounts* **97**, 240 (1997).
(54) A. Komornicki and H. F. King, *Journal of Chemical Physics* **134**, 244115 (2011).
(55) E. G. Hohenstein, R. M. Parrish, and T. J. Martinez, *Journal of Chemical Physics* **137**, 044103 (2012).
(56) E. G. Hohenstein, R. M. Parrish, C. D. Sherrill, and T. J. Martinez, *Journal of Chemical Physics* **137**, 221101 (2012).
(57) R. M. Parrish, E. G. Hohenstein, T. J. Martinez, and C. D. Sherrill, *Journal of Chemical Physics* **137**, 224106 (2012).

Chapter 7

Reduced Explicitly Correlated Hartree-Fock Approach within the Nuclear-Electronic Orbital Framework: Applications to Positronic Molecular Systems[‡]

7.1. Introduction

Positron-bound species play an important role in many fields, such as astrophysics,[1] chemistry,[2] and solid state physics.[3,4] In addition, positron emission tomography has been established as a useful imaging tool at the molecular level for medical applications.[5] In these various contexts, the main property of interest is the annihilation between electrons and positrons. The dominant annihilation mechanism yields two photons, and its rate, $\Gamma_{2\gamma}$, is a quantity that is both experimentally and theoretically accessible.[2] Furthermore, theoretical methods can be used to calculate the electronic and positronic densities, thereby providing insight into the locations and relative magnitudes of positron annihilation.

A variety of computational methods have been developed to investigate positron-bound species. Previously we studied several positronic molecular systems within the nuclear-electronic orbital (NEO) framework using the Hartree-Fock (NEO-HF), full configuration interaction (NEO-FCI), second-order perturbation theory (NEO-MP2), and explicitly correlated (NEO-XCHF) approaches.[6-8] In the application of NEO to positronic systems, all electrons and the

[‡]Reproduced with permission from:
A. Sirjoosingh, M. V. Pak, C. Swalina, and S. Hammes-Schiffer, "Reduced explicitly correlated Hartree-Fock approach within the nuclear-electronic orbital framework: Applications to positronic molecular systems," *J. Chem. Phys.* **139**, 034103 (2013). © 2013 American Institute of Physics

positron are treated quantum mechanically, and all nuclei are treated classically. The NEO-XCHF approach includes electron-positron correlation using Gaussian-type geminal functions within the variational self-consistent-field procedure. These positronic systems have also been studied using highly accurate but more computationally expensive explicitly correlated methods, including the previously denoted explicitly correlated Gaussian (ECG)[9-15] method and the stochastic variational method (SVM).[16-19] The NEO-XCHF approach is more computationally practical than the ECG and SVM approaches for two reasons. First, the NEO-XCHF approach treats only electron-positron correlation explicitly, while the ECG and SVM approaches treat both electron-electron and electron-positron correlation explicitly, resulting in more complicated wavefunctions. As discussed previously,[7,8] electron-positron dynamical correlation is more significant than electron-electron dynamical correlation because of the attractive electron-positron Coulomb interaction. Second, the ECG and SVM approaches typically require thousands of parameters that are optimized with a stochastic variational scheme, whereas the NEO-XCHF approach introduces a relatively small number of parameters that are optimized through iterative schemes. Several other methods have also been developed and applied to positron-containing species.[20-31]

The availability of ECG and SVM results for small positron-containing species provides reliable benchmarks for testing the much more computationally tractable NEO methods. In a previous study, we performed benchmarking studies for several positronic systems.[8] The NEO-HF method provides qualitative insight into the location of electron-positron annihilation in certain cases but is generally unreliable. The NEO-XCHF method, however, provides qualitatively reasonable electronic and positronic densities and semi-quantitative annihilation rates and related quantities. The success of the NEO-XCHF method relies on the property that

annihilation rates depend on regions of small electron-positron distances rather than on the quality of the overall electron-positron wavefunction. Thus, a suitable treatment of the essential electron-positron dynamical correlation leads to accurate annihilation rates.

Recently we developed an alternative reduced version of the explicitly correlated wavefunction approach within the NEO framework, denoted NEO-RXCHF. The fundamental difference between the original XCHF approach and the more recently developed RXCHF approach is that only a subset of the electronic orbitals is explicitly correlated to the orbital of the quantum nucleus (or positron) in the RXCHF approach. The theory underlying the RXCHF approach is presented in the accompanying paper.[32] The RXCHF approach is well-suited to the study of positron-containing species, especially those with significant positronium (i.e., an electron-positron pair) character. By explicitly correlating only one electronic orbital to the positronic orbital in the RXCHF approach, the orbital product can be interpreted as a description of a single positronium particle. The total electron-positron wavefunction is optimized to accurately describe the essential electron-positron correlation within this positronium particle. In contrast, the XCHF approach explicitly correlates all electrons to the positron in the same manner, often leading to a less accurate description of the most important short-ranged electron-positron interaction because the geminal functions are used to account for other interactions. Thus, properties such as the electron-positron contact density, as well as annihilation rates, can be captured much more effectively with the RXCHF approach than with the XCHF approach.

In this study, we apply NEO-RXCHF to several positronic systems to demonstrate its improvement over NEO-XCHF. The results clearly illustrate that explicitly correlating a single electronic orbital to the positronic orbital offers significant advantages in accuracy over correlating all of the electronic orbitals to the positronic orbital within this framework. We also

compare our calculations to the highly accurate ECG and SVM data available in the literature[12,15-17,19] and show that the NEO-RXCHF approach provides almost as accurate annihilation rates as these more sophisticated methods, while offering significant advantages in computational tractability. Furthermore, we evaluate two different approximations within the RXCHF approach related to the electronic exchange contributions and provide evidence that these approximations exert only negligible effects on the accuracy at a substantially lower computational cost.

This paper is organized as follows. In Section 7.2, we briefly summarize the theory presented in detail in the accompanying paper.[32] We also discuss the calculation of relevant expectation values and define the densities that will be analyzed in this paper, as well as describe the schemes to determine the Gaussian-type geminal function parameters. In Section 7.3, we examine three positron-containing species: positron-lithium, lithium positride, and positron-lithium hydride. In the final section, we present concluding remarks and discuss future directions.

7.2. Theory

A detailed account of the RXCHF method and related approximations is provided in the accompanying paper.[32] In this section, we repeat the key definitions to facilitate a comparison among the various methods in the subsequent sections. Then we define the quantities necessary for the calculation of densities and annihilation rates. The formalism below is presented for the case of one positron and N electrons, with only one electronic orbital geminal-coupled (i.e., explicitly correlated) to the positronic orbital.

A. Wavefunction definitions

The fully antisymmetric RXCHF wavefunction, denoted RXCHF-fe, is given by

$$\Psi^{\text{RXCHF-fe}}\left(\mathbf{x}^e,\mathbf{x}^p\right)=\frac{\chi^p\left(\mathbf{x}^p\right)}{\sqrt{N!}}\begin{vmatrix}\chi_1^e\left(\mathbf{x}_1^e\right)g\left(\mathbf{r}_1^e,\mathbf{r}^p\right) & \chi_2^e\left(\mathbf{x}_1^e\right) & \cdots & \chi_N^e\left(\mathbf{x}_1^e\right)\\ \chi_1^e\left(\mathbf{x}_2^e\right)g\left(\mathbf{r}_2^e,\mathbf{r}^p\right) & \chi_2^e\left(\mathbf{x}_2^e\right) & \cdots & \chi_N^e\left(\mathbf{x}_2^e\right)\\ \vdots & \vdots & \ddots & \vdots\\ \chi_1^e\left(\mathbf{x}_N^e\right)g\left(\mathbf{r}_N^e,\mathbf{r}^p\right) & \chi_2^e\left(\mathbf{x}_N^e\right) & \cdots & \chi_N^e\left(\mathbf{x}_N^e\right)\end{vmatrix}, \tag{7.1}$$

where χ^p is the positronic spin orbital, $\left\{\chi_a^e\right\}$ are the electronic spin orbitals, and $\mathbf{x}^p$ and $\left\{\mathbf{x}_i^e\right\}$ are the positronic and electronic coordinates, respectively. The Gaussian-type geminal (GTG) functions are defined as

$$g\left(\mathbf{r}_i^e,\mathbf{r}^p\right)=\sum_{k=1}^{N_{\text{gem}}}b_k e^{-\gamma_k\left|\mathbf{r}_i^e-\mathbf{r}^p\right|^2}. \tag{7.2}$$

In Eq. (7.1), one electronic orbital, χ_1^e, is coupled to the positronic orbital via the GTGs. We refer to this orbital as the geminal-coupled electronic orbital.

An approximation to the fully antisymmetric RXCHF wavefunction in Eq. (7.1) is given by

$$\Psi^{\text{RXCHF-ne}}\left(\mathbf{x}^e,\mathbf{x}^p\right)=\frac{\chi^p\left(\mathbf{x}^p\right)\chi_1^e\left(\mathbf{x}_1^e\right)g\left(\mathbf{r}_1^e,\mathbf{r}^p\right)}{\sqrt{(N-1)!}}\begin{vmatrix}\chi_2^e\left(\mathbf{x}_2^e\right) & \cdots & \chi_N^e\left(\mathbf{x}_2^e\right)\\ \vdots & \ddots & \vdots\\ \chi_2^e\left(\mathbf{x}_N^e\right) & \cdots & \chi_N^e\left(\mathbf{x}_N^e\right)\end{vmatrix}. \tag{7.3}$$

In this ansatz, the geminal-coupled electronic orbital is not in the Slater determinant, so the overall wavefunction is not antisymmetric with respect to exchange of all electronic coordinates. In other words, this ansatz identifies one electron, with coordinate $\mathbf{x}_1^e$, to be distinguishable from the others. We denote this electron as the "special" electron and the other electrons as the "regular" electrons.

As defined above, the RXCHF-fe method contains full exchange contributions between the geminal-coupled electronic orbital and the regular electronic orbitals, while the RXCHF-ne method neglects these exchange contributions. In the accompanying paper,[32] we also describe an intermediate approximation, RXCHF-ae, which includes an additional approximate exchange contribution to the RXCHF-ne energy expression. The energy expression that is optimized variationally with respect to the orbital coefficients for the RXCHF-ae method is

$$E^{\text{RXCHF-ae}} = E^{\text{RXCHF-ne}} + E_{\text{ex}}, \tag{7.4}$$

where $E^{\text{RXCHF-ne}}$ is the expectation value of the total system Hamiltonian with the wavefunction given in Eq. (7.3), and E_{ex} is an approximate exchange term that also depends on the orbitals.[32] In the remainder of this paper, we report results for all three of these methods to investigate the validity of the RXCHF-ne and RXCHF-ae approximations to RXCHF-fe.

For comparative purposes, we include the definition of the NEO-XCHF wavefunction,

$$\Psi^{\text{XCHF}}\left(\mathbf{x}^e,\mathbf{x}^p\right) = \Phi^e\left(\mathbf{x}^e\right)\chi^p\left(\mathbf{x}^p\right)\left[1+G\left(\mathbf{r}^e,\mathbf{r}^p\right)\right], \tag{7.5}$$

where $\Phi^e\left(\mathbf{x}^e\right) = \left|\chi_1^e\left(\mathbf{x}_1^e\right)\cdots\chi_N\left(\mathbf{x}_N^e\right)\right\rangle$ is a Slater determinant of N electronic spin orbitals and $G\left(\mathbf{r}^e,\mathbf{r}^p\right) = \sum_{i=1}^{N} g\left(\mathbf{r}_i^e,\mathbf{r}^p\right)$. The NEO-XCHF wavefunction explicitly correlates all electronic orbitals to the positronic orbital via the same GTGs.

In the XCHF wavefunction given by Eq. (7.5), the geminal part of the wavefunction is $1+G$. The physical basis of this choice is to ensure that at large electron-positron distances (i.e., where G vanishes), the XCHF wavefunction approaches the NEO-HF wavefunction (i.e., a product of electronic and positronic Slater determinants). In the RXCHF wavefunctions given by Eqs. (7.1) and (7.3), the geminal part of each wavefunction is G. The physical basis of this choice is to ensure that at large electron-positron distances, the entire RXCHF wavefunction

vanishes. The choice of this ansatz is appropriate for the positronic systems studied herein, as the electron-positron pair described by the geminal-coupled orbitals could be interpreted as a single positronium particle. In this case, when the geminal-coupled electron-positron pair is well-separated, the wavefunction vanishes due to the nonexistence of positronium.

It is straightforward to interconvert between the $1+G$ and G forms of the geminal part of the XCHF and RXCHF wavefunctions. To convert the geminal part of the XCHF wavefunction in Eq. (7.5) to the G form, an additional geminal with $b=-1/N$ and $\gamma \approx 0$ can be included. To convert the geminal part of the RXCHF wavefunctions in Eqs. (7.1) and (7.3) to the $1+G$ form, an additional geminal with $b=1$ and $\gamma \approx 0$ can be included. The latter procedure will likely be important for studying systems with quantum nuclei, such as protons, with the RXCHF method.

In the calculations described below, we utilize a fixed set of geminal parameters and variationally optimize the energy with respect to the orbital coefficients. The energy expressions for the NEO-RXCHF-fe, NEO-RXCHF-ne, and NEO-XCHF methods are obtained from the expectation value of the total system Hamiltonian with the associated wavefunctions given in Eqs. (7.1), (7.3), and (7.5), respectively. These methods are fully variational with respect to the orbital coefficients for the specified energy expression with a fixed set of geminal parameters. The energy expression for the NEO-RXCHF-ae method is given by Eq. (7.4), and this energy, including the approximate exchange term E_{ex}, is minimized with respect to the orbital coefficients. The NEO-RXCHF-ae method is not subject to the variational theorem, however, because there is no wavefunction that directly leads to this energy expression.

B. Expectation values and annihilation rates

Expectation values are calculated using the standard definition

$$\langle \mathrm{A} \rangle = \frac{\left\langle \Psi \middle| \hat{\mathrm{A}} \middle| \Psi \right\rangle}{\langle \Psi | \Psi \rangle}, \tag{7.6}$$

where $\hat{\mathrm{A}}$ is the operator corresponding to the observable $\langle \mathrm{A} \rangle$. For the calculations presented in this paper, we use $\Psi = \Psi^{\text{RXCHF-fe}}$ to evaluate expectation values for the fully antisymmetric ansatz and use $\Psi = \Psi^{\text{RXCHF-ne}}$ for the approximate ansatz with one distinguishable electron. The RXCHF-ae method has no corresponding wavefunction, as it involves the addition of an approximate exchange contribution to the energy, so we use $\Psi = \Psi^{\text{RXCHF-ne}}$ to calculate expectation values for the RXCHF-ae method, noting that the eigenvectors used in the expansion of Eq. (7.6) correspond to a state that minimizes the energy expression including the approximate exchange contribution.

The expectation values of interest in the study of positronic systems are related to positronic and electronic density expressions. The single-particle positronic density is defined as

$$\rho^{p}(\mathbf{r}) = \left\langle \delta\left(\mathbf{r}^{p} - \mathbf{r}\right) \right\rangle = \frac{\left\langle \Psi \middle| \delta\left(\mathbf{r}^{p} - \mathbf{r}\right) \middle| \Psi \right\rangle}{\langle \Psi | \Psi \rangle}, \tag{7.7}$$

and the single-particle electronic density is defined as

$$\rho^{e}(\mathbf{r}) = \left\langle \sum_{i=1}^{N} \delta\left(\mathbf{r}_{i}^{e} - \mathbf{r}\right) \right\rangle = \frac{\left\langle \Psi \middle| \sum_{i=1}^{N} \delta\left(\mathbf{r}_{i}^{e} - \mathbf{r}\right) \middle| \Psi \right\rangle}{\langle \Psi | \Psi \rangle}. \tag{7.8}$$

The electron-positron contact density is defined as

$$\rho^{ep}(\mathbf{r}) = \left\langle \sum_{i=1}^{N} \delta\left(\mathbf{r}_{i}^{e} - \mathbf{r}\right) \delta\left(\mathbf{r}^{p} - \mathbf{r}\right) \right\rangle, \tag{7.9}$$

which enables the calculation of the average contact density,

$$\left\langle \delta_{ep} \right\rangle = \left\langle \sum_{i=1}^{N} \delta\left(\mathbf{r}^{p} - \mathbf{r}_{i}^{e}\right) \right\rangle = \int \mathrm{d}\mathbf{r} \rho^{ep}\left(\mathbf{r}\right). \tag{7.10}$$

The two-photon annihilation rate is defined as

$$\Gamma_{2\gamma} = 4\pi r_e^2 c \left\langle \sum_{i=1}^{N} \delta\left(\mathbf{r}^{p} - \mathbf{r}_{i}^{e}\right) \hat{O}_{e_i p}^{s} \right\rangle = 4\pi r_e^2 c \frac{\left\langle \Psi \left| \sum_{i=1}^{N} \delta\left(\mathbf{r}^{p} - \mathbf{r}_{i}^{e}\right) \hat{O}_{e_i p}^{s} \right| \Psi \right\rangle}{\left\langle \Psi | \Psi \right\rangle}, \tag{7.11}$$

where r_e is the classical electron radius, c is the speed of light, and $\hat{O}_{e_i p}^{s}$ is the spin projection operator for the i^{th} electronic coordinate. The action of the spin projection operator $\hat{O}_{e_i p}^{s}$ on Ψ yields a nonzero result only when the spin of the i^{th} electron is opposite the spin of the positron (i.e., singlet coupled). This rate is calculated by spin-projecting the appropriate expectation value expression given on the right-hand side of Eq. (7.11). Further details of this projection are given in Section 7.3.

C. Geminal function parameters

Both the XCHF and RXCHF methods require the determination of the GTG function parameters, $\left\{\left(b_k, \gamma_k\right)\right\}$. In previous work on positronic systems,[8] the geminal parameters used for the NEO-XCHF wavefunction were fit variationally for the positronium hydride (PsH) system. In this work, $N_{\text{gem}} = 8$ GTGs were determined to be sufficient to converge the densities and annihilation rates for PsH. To study other positronic systems, the parameters $\left\{b_k\right\}$ from the PsH system were scaled by a single factor to account for differences in the relative amount of explicit

correlation versus Hartree-Fock character in the XCHF wavefunction. This procedure corresponds to expressing the NEO-XCHF wavefunction in Eq. (7.5) as

$$\Psi^{\mathrm{XCHF}}\left(\mathbf{x}^e,\mathbf{x}^p\right)=\Phi^e\left(\mathbf{x}^e\right)\chi^p\left(\mathbf{x}^p\right)\left[1+\beta G^{\mathrm{PsH}}\left(\mathbf{r}^e,\mathbf{r}^p\right)\right]. \tag{7.12}$$

The scaling factor β was determined variationally for each different system.

We have applied two different schemes for determining the geminal parameters for the XCHF and RXCHF wavefunctions:

(i) We employ the $1+G$ form for the geminal part of the wavefunction. To determine specific geminal function parameters for each system, we start with the geminal parameters for $N_{\mathrm{gem}}=8$ from Ref. 8 (i.e., optimized for the NEO-XCHF wavefunction for the PsH system) and optimize the scaling factor β in Eq. (7.12) at the appropriate level of theory (i.e., XCHF or RXCHF) for each different system.

(ii) We employ the G form for the geminal part of the wavefunction and determine a fixed set of parameters used for all systems. For this purpose, we variationally optimized the geminal parameters with $N_{\mathrm{gem}}=8$ for a model one-electron-one-positron e^-e^+X system, where X is a fixed classical point charge of +0.5. This optimization was performed for an ansatz given by $\Psi\left(\mathbf{x}_1^e,\mathbf{x}^p\right)=\chi_1^e\left(\mathbf{x}_1^e\right)\chi^p\left(\mathbf{x}^p\right)g\left(\mathbf{r}_1^e,\mathbf{r}^p\right)$ because the RXCHF and XCHF wavefunctions coincide for a one-electron-one-positron system. After optimization of the geminal parameters for this model system, the resulting GTG parameters, which are given in Table 7.1, were used for all other systems studied.

The second scheme relies on the transferability of the GTG parameters calculated for the e^-e^+X model system, whereas the first scheme requires optimization of the scaling factor β for each new system studied. For positronic systems, a fixed set of GTG parameters given by scheme (ii)

is likely to be suitable for different systems because the short-ranged electron-positron interactions within a single positronium are expected to remain relatively invariant to the surrounding chemical environment.

In this paper, we examine three systems: positron-lithium (e^+Li), lithium positride (LiPs), and positron-lithium hydride (e^+LiH). For these three systems, GTG scheme (ii) resulted in significantly more accurate annihilation rates than GTG scheme (i). The physical basis for the advantages of GTG scheme (ii) is the positronium character inherent to these types of positron-bound molecular systems. As discussed at the end of Section 7.2A, the *G* form for the geminal part of the wavefunction is more suitable for these types of systems. For this reason, we present only the results for GTG scheme (ii) in the main paper and provide the results for GTG scheme (i) in Appendix F. We also applied the RXCHF method to the PsH system. Because this system has only two electrons and has limited positronium character, the RXCHF method does not provide significant advantages over the XCHF method, and GTG scheme (i) is more suitable for this specific system. For completeness, the results for PsH are also provided in Appendix F, although the XCHF method has already been shown to be sufficient.

7.3. Results and Discussion

In this section, we compare XCHF, RXCHF-fe, RXCHF-ne, and RXCHF-ae results to benchmark ECG and SVM data. We present results for a three-electron system, e^+Li, as well as two four-electron systems, LiPs and e^+LiH. In each subsection, we briefly discuss specific issues arising for the calculation of annihilation rates at the RXCHF level, and then we compare the single particle and contact densities, as well as the annihilation rates and average contact densities.

A. Positron-lithium

The e^+Li system consists of three electrons, one positron, and a lithium nucleus. It has not been studied previously with NEO-XCHF but has been studied extensively with the SVM approaches. This system provides the best opportunity to test the RXCHF-ne and RXCHF-ae approximations by comparison to the fully antisymmetric RXCHF-fe method, as the formulation for the latter presented in the accompanying paper[32] is only valid for up to three electrons.

The Gaussian basis set used for this system consists of 10 *s* functions for the electronic orbitals and 8 *s* functions for the positronic orbital. This basis set was constructed from the 8s/6s basis set from Ref. 8 used for the LiPs system by adding two diffuse functions to each basis set using the same even-tempered scheme described in that study. We determined that additional diffuse basis functions were required in order to capture the loosely bound positronium character and that the 10s/8s basis set was well-converged with respect to adding more diffuse functions. (See Table F.3 for values obtained using a 12s/10s basis set.)

The annihilation rate defined in Eq. (7.11) can be determined for two cases: a singlet case, in which the unpaired electron has the opposite spin as the positron, and a triplet case, in which the unpaired electron has the same spin as the positron. Without loss of generality, we assume the spin of the unpaired electron to be α. In the singlet-coupled (triplet-coupled) case, the positron has a spin of β (α), and Eq. (7.11) is evaluated using appropriate action of the spin-projection operator.

In the singlet-coupled case, the annihilation rate is defined as

$$\Gamma_s \equiv 4\pi r_e^2 c \left\langle \delta_{ep} \right\rangle_{\text{sing}}, \tag{7.13}$$

where $\left\langle \delta_{ep} \right\rangle_{\text{sing}}$ is obtained from the right-hand side of Eq. (7.11) after action of the spin projection operator for the singlet-coupled system. This procedure involves preserving terms in the contact density coupling the positron and the unpaired electron and one core electron. In the triplet-coupled case, the annihilation rate is defined as

$$\Gamma_t \equiv 4\pi r_e^2 c \left\langle \delta_{ep} \right\rangle_{\text{trip}}, \tag{7.14}$$

where $\left\langle \delta_{ep} \right\rangle_{\text{trip}}$ is obtained from the right-hand side of Eq. (7.11) after action of the spin projection operator for the triplet-coupled system. This procedure involves preserving terms in the contact density coupling the positron and one core electron. Finally, the spin-averaged annihilation rate, which weights the above two contributions by their associated probabilities, is defined as

$$\Gamma_a = \frac{\Gamma_s + 3\Gamma_t}{4}. \tag{7.15}$$

Additional information about the spin-projection procedure, including expressions for the singlet- and triplet-coupled cases, is given in Section F.2.

The energies, average contact densities, and annihilation rates for e^+Li are reported in Table 7.2. For reasons discussed above, GTG scheme (ii), which assumes the *G* form for the geminal part of the wavefunction, is expected to provide physically meaningful results for most positronic systems. For this system and the next two systems, we only present the results obtained with this GTG scheme. The results obtained with GTG scheme (i) are less accurate and are included in Appendix F.

For the e^+Li system, the RXCHF energies are lower than the XCHF energy, but all of these NEO energies are significantly higher than the SVM energy. As mentioned above, the total energy is not expected to be as accurate for the RXCHF and XCHF methods as for the SVM

due to the neglect of electron-electron correlation, as well as significantly fewer optimized parameters, but the objective of this study is to focus on the short-ranged electron-positron properties, such as the annihilation rates. The average contact densities and annihilation rates for all methods are in very good agreement with the benchmark SVM data. In particular, the XCHF calculations seem to be very accurate, but as discussed below, this agreement is likely fortuitous. The RXCHF methods are able to predict reasonable triplet annihilation rates, with RXCHF-ae and RXCHF-fe slightly more accurate than RXCHF-ne. Since triplet annihilation for e^+Li quantifies annihilation between the positron and a single core electron, we expect that exchange effects between core and valence electrons should impact this property. The triplet rate predicted by XCHF is very poor, possibly due to the incorrect description of the core electron-positron correlation, which is required to have the same form as the valence electron-positron correlation within the XCHF ansatz. We emphasize that the computational cost of the RXCHF methods is substantially lower than that of the SVM approach. Thus, this level of agreement in the annihilation rates is considered to be promising for future applications of RXCHF to larger chemical systems.

The electronic and positronic single-particle densities, as well as the contact densities, for e^+Li calculated with the XCHF and RXCHF-fe methods are depicted in Figure 7.1a and 7.1b, respectively. The RXCHF-fe method leads to accurate single-particle densities, but the XCHF method does not lead to even qualitatively reasonable single-particle densities, as illustrated by the deviations from the SVM data (dashed lines). The first peak in the electronic density (blue curve) corresponds to the core electrons. The broader peak, which exhibits coalescence of the electronic and positronic densities at large distances (red and blue curves), is typical of weakly bound positronium. On the basis of the single-particle densities in Figure 7.1a, the positronium

appears to dissociate for the XCHF method but is weakly bound for the RXCHF-fe method, as is the case with the SVM. Moreover, the XCHF contact density (green curve) in Figure 7.1a is shifted to much greater distances, further indicative of positronium dissociation. In contrast, the contact density in Figure 7.1b clearly demonstrates that RXCHF-fe predicts a loosely-bound positronium system. In particular, the smaller peak at very short distances indicates coalescence of the positron with the core electrons, while the larger peak indicates coalescence of the positron with the valence electron. Given the qualitatively incorrect single-particle and contact densities generated by the XCHF method, the agreement between the XCHF and SVM two-photon annihilation rates, as given in Table 7.2, is presumed to be fortuitous.

In contrast, the RXCHF approach provides reasonably accurate single-particle densities, as well as annihilation rates. Figure 7.2 depicts a comparison of the single-particle and contact densities for the three different RXCHF methods, which differ in the treatment of electronic exchange. Note that all three methods lead to very similar densities. In particular, the RXCHF-ae and RXCHF-fe methods lead to virtually identical single-particle and contact densities. We also analyzed the relative contribution of the approximate exchange term, E_{ex}, in Eq. (7.4). In the RXCHF-ae method, the total energy given in Eq. (7.4), including the approximate exchange term E_{ex}, is minimized with respect to the orbital coefficients. Thus, the RXCHF-ae energy differs from the RXCHF-ne energy not only because of the approximate exchange term, but also because of the difference in orbital coefficients obtained from the two distinct minimization procedures. To facilitate the quantification of E_{ex}, we analyzed the contribution of the exchange term for a fixed set of orbital coefficients. Using the orbital coefficients obtained from the RXCHF-fe calculation, we found that E_{ex} accounts for 99.98% of the difference between the RXCHF-fe and RXCHF-ne energies. This result is consistent with the observation that the

independent optimizations of the RXCHF-fe and RXCHF-ae energies resulted in nearly identical solutions, as demonstrated by the virtually indistinguishable densities.

For the e^+Li system, the RXCHF method exhibits a clear advantage over the XCHF method. The RXCHF single-particle densities, as well as the annihilation rates, are in very good agreement with the benchmark SVM results. Furthermore, the RXCHF-ne and RXCHF-ae methods provide reasonable approximations to the fully antisymmetric RXCHF-fe method, while offering significant advantages with respect to computational tractability, as will be discussed in Section 7.4. In the next two subsections, we proceed to study two four-electron systems and perform only RXCHF-ne and RXCHF-ae calculations due to limitations of the RXCHF-fe formalism.

B. Lithium positride

The LiPs system consists of four electrons, one positron, and a lithium nucleus. It has been studied previously with NEO-XCHF[8] as well as ECG methods. As mentioned above, we cannot study this system at the RXCHF-fe level due to limitations of the formulation presented in the accompanying paper,[32] but we can still compare the RXCHF-ne and RXCHF-ae methods to the benchmark ECG data. The Gaussian basis set used for this system was the 8s/6s basis set obtained from Ref. 8, where it was optimized for XCHF calculations on LiPs.

The annihilation rate defined in Eq. (7.11) is easily calculated for the XCHF method because LiPs is treated in an RHF formalism, where orbitals are doubly occupied by electrons of opposite spin. In this case, the spin projection operator picks out exactly half of the unprojected $\langle \delta_{ep} \rangle$. In the RXCHF-ae and RXCHF-ne methods, however, LiPs is described by a UHF formalism, where the degeneracy of one of the electron pairs is removed because one of the electrons is distinguished from the others, and the three remaining electrons are described with

an unrestricted approach (i.e., two different modified Hartree-Fock equations for α and β regular electronic orbitals). In this framework, two different scenarios can lead to annihilation. Without loss of generality, we assume the spin of the positron to be α. The two cases are as follows: (1) the distinguishable electron is of spin β, and there are two uncoupled electrons of spin α and one of spin β; and (2) the distinguishable electron is of spin α, and there are two uncoupled electrons of spin β and one of spin α. In the first case, the positron can annihilate with the distinguishable electron and the one remaining electron with spin β. In the second case, the positron can annihilate only with the two uncoupled electrons of spin β. As there is an equal probability of either of these scenarios occurring, we calculate the rate corresponding to each of these scenarios and report the average of their values. Note that in the limit that the spatial orbital for the distinguishable electron is identical to that of the remaining valence electron (i.e., a closed-shell scenario), the annihilation rate is calculated as in XCHF, i.e., it is proportional to half of the unprojected $\langle \delta_{ep} \rangle$.

The energies, average contact densities, and annihilation rates for LiPs are reported in Table 7.3. For this system, the RXCHF-ne and RXCHF-ae energies are lower than the XCHF energy, and the average contact densities and annihilation rates for the RXCHF methods are significantly better than the corresponding XCHF values. Furthermore, the RXCHF-ne and RXCHF-ae methods lead to very similar results. We found that the RXCHF annihilation rates calculated for the two scenarios discussed above are very different, but their average value matches the benchmark values well. For example, the RXCHF-ne rate of 1.9405 ns^{-1} is the average value of 3.8245 and 0.056534 ns^{-1}. We include these quantities for this representative example to demonstrate quantitatively that we are far from the situation assumed by XCHF (i.e., two doubly occupied orbitals). The main contribution to the rate arises from the first scenario, in

which the distinguishable electron can annihilate with the positron. The single-particle densities for XCHF and RXCHF-ae are depicted in Figure 7.3 and are both very similar to the benchmark SVM plots. The RXCHF-ne densities are virtually indistinguishable from the RXCHF-ae densities. As discussed in Section 7.4, the RXCHF-ae and RXCHF-ne methods offer substantial computational savings over the XCHF method for this system.

Although not reported here, we also performed RXCHF calculations on LiPs with more diffuse basis sets. We followed the procedure described for the e^+Li system, where two diffuse basis functions were added to the 8s/6s basis set obtained from Ref. 8. With this increased flexibility, the RXCHF calculations seemed to promote positronium dissociation with both GTG schemes. The enhanced diffusivity of this basis set is unsuitable for this study in that we aim to describe the bound state of LiPs and not its dissociated state. Thus, careful attention should be directed toward the choice of basis sets when studying these types of weakly bound systems.

C. Positron lithium hydride

The e^+LiH system consists of four electrons, one positron, a lithium nucleus, and a proton. It has been studied previously with XCHF[8] as well as ECG methods. We adopt the model described in Ref. 8, where the distance between the lithium nucleus and the proton is fixed to its equilibrium distance of 3.015 a.u. The annihilation rates for the e^+LiH system are calculated in the same way as for the LiPs system, and the same averaging procedure for the RXCHF rates was used.

The Gaussian basis set used for this system was obtained from Ref. 8, where it was determined using a detailed procedure involving both NEO-HF and XCHF optimizations. The complication of this basis set optimization procedure is due in part to the computational expense incurred when performing XCHF calculations on systems of this size. The electronic basis set is

the pc0 basis set of Jensen[33,34] placed on the lithium and hydrogen centers, and the positronic basis set consists of 6 *s* functions. In this optimization procedure, the positions of the positronic basis functions are determined at the NEO-HF level, and the exponents of these basis functions are determined using XCHF with $N_{\text{gem}} = 2$. We refer to this basis set, obtained directly from Ref. 8, as pc0/6s.

Due to the decreased computational cost of RXCHF calculations for this system, additional basis set optimizations could be performed at the RXCHF level. Thus, we present results for the RXCHF methods where both the positions and the exponents of the positronic basis functions were optimized at the RXCHF-ae level using GTG scheme (ii) with the pc0 basis set. Optimizations were also performed at the RXCHF-ne level, but the RXCHF-ae basis set gave improved results. In addition, we were able to increase the level of the electronic basis set to pc1[33,34] for the RXCHF calculations, although the positronic basis function centers were not reoptimized at this level. We denote this basis set as pc1/6s-opt.

The energies, average contact densities, and annihilation rates for e^+LiH are reported in Table 7.4. First we analyze the results with the basis set pc0/6s to allow a comparison to the XCHF results. Although the energies obtained with the RXCHF methods are relatively high, the average contact densities and annihilation rates are considerably better than the XCHF values. As mentioned above, our objective is to obtain accurate annihilation rates and other short-ranged electron-positron properties with a computationally tractable method. In this case, the short-ranged electron-positron interaction is optimized to the detriment of other parts of the wavefunction, thereby leading to higher total energies.

The single-particle electronic (blue), positronic (red), and contact (green) densities are depicted for the XCHF and RXCHF-ae methods in Figure 7.4a and b, respectively. The

analogous RXCHF-ne densities are very similar to the RXCHF-ae densities for this basis set. Figure 7.4 illustrates that these slices of the XCHF and RXCHF-ae densities are qualitatively similar. The increased quantitative accuracy of RXCHF-ae compared to XCHF for the average contact density and annihilation rate demonstrates, however, that the RXCHF methods yield a more accurate description of the overall electron-positron interaction. In comparison to the ECG method, the RXCHF-ae method leads to exaggerated contact density (green curves) on the left peak (centered on Li), does not capture the secondary peak in the electronic densities (blue curves) very well, and does not capture the shoulder in the positronic peak (red curves). In an effort to improve the densities, we took advantage of the computational tractability of the RXCHF methods and reoptimized the positronic basis set.

The RXCHF results with the pc1/6s-opt basis set are labeled with asterisks in Table 7.4. As discussed above, the positronic basis set was optimized at the RXCHF-ae level while using the pc0 basis set, and the electronic basis set was increased to pc1 for the subsequent calculations. We did not perform XCHF calculations with this basis set due to the high computational costs. The energies, average contact densities, and rates are similar to those calculated with the pc0/6s basis set and are in very good agreement with the ECG results.

The densities obtained with the RXCHF-ae method and the pc1/6s-opt basis set are depicted in Figure 7.4c. The corresponding plots for the RXCHF-ne method are very similar and are provided in Figure F.5. As a result of basis set optimization, we observe a clear improvement in the single-particle and contact densities. In particular, RXCHF is able to capture the shoulder in the positronic density (red curves). This shoulder was not evident when the positronic basis set optimization was carried out at the RXCHF-ne level, indicating that the inclusion of approximate exchange offers some advantage in accuracy. In addition, RXCHF in conjunction with the pc1

electronic basis set is able to describe the secondary peak in the electronic single-particle density (blue curves) more accurately. The contact density (green curves) appears to become less accurate near the right peak localized around H but improves significantly near the left peak localized around Li. Further refinements of both basis sets at the RXCHF-ae level could possibly improve these densities, although the average contact densities and rates evaluated over the entire wavefunction are already very accurate.

These results for e^+LiH indicate that both the RXCHF-ne and RXCHF-ae methods perform quantitatively well with respect to average contact densities and annihilation rates, and both RXCHF methods outperform XCHF. The increased tractability of the RXCHF methods over the XCHF method allows more consistent optimizations of the basis sets used for the e^+LiH system. Although the contact density along the Li-H axis exhibits some deviations from the ECG contact density, the RXCHF wavefunction appears to be more accurate than the XCHF wavefunction, and further refinement of the basis sets is expected to provide additional improvement. Finally, we observed a difference between the RXCHF-ne and RXCHF-ae methods in that the latter is able to capture the shoulder in the positronic density, which is not captured at all by XCHF or when positronic basis set optimizations were performed at the RXCHF-ne level. Thus, the inclusion of approximate exchange appears to offer some advantage over the neglect of exchange between the geminal-coupled electron and the regular electrons at comparable computational expense.

7.4. Concluding Remarks

In this paper, we described the application of the recently developed RXCHF methods to several positron-containing species: e^+Li, LiPs, and e^+LiH. The RXCHF results were compared

to highly accurate SVM and ECG data for benchmarking purposes. For all three of these systems, the RXCHF methods provide reasonably accurate two-photon annihilation rates, average contact densities, electronic and positronic single-particle densities, and electron-positron contact densities. Moreover, the RXCHF method was shown to be significantly more accurate than the XCHF method in reproducing these properties.

The reason for the superiority of RXCHF over XCHF lies in the underlying assumptions of the two methods. In both cases, the GTGs can be chosen to accurately represent the short-ranged electron-positron dynamical correlation by optimizing the geminal parameters for a one-electron-one-positron model system. According to the XCHF ansatz for the wavefunction, however, all electronic orbitals are explicitly correlated to the positronic orbital using the same GTGs (i.e., the same geminal parameters). For positronic molecular systems, the explicit correlation of the positron to the core and valence electronic orbitals in the same manner can lead to electron-positron wavefunctions in which the geminal functions are used to account for interactions other than the short-ranged electron-positron dynamical correlation. During the variational optimization of the electronic and positronic orbitals, the XCHF wavefunction is not optimized to accurately describe the specific short-ranged electron-positron interaction but rather is more globally optimized.

In contrast, the RXCHF method correlates only a single electronic orbital with the positronic orbital, and the geminal functions are used mainly to account for the short-ranged electron-positron dynamical correlation. Thus, the RXCHF wavefunction is optimized to produce a highly accurate description of the short-ranged electron-positron interaction, although the overall wavefunction may not be optimal. For these reasons, the RXCHF method leads to more accurate values of the annihilation rates and other local properties that depend strongly on

the accuracy of the electronic-positronic wavefunction in the region of short electron-positron distances.

The investigation of the three-electron e^+Li system allowed us to compare the fully antisymmetric RXCHF-fe method to the RXCHF-ne and RXCHF-ae methods, which either neglect or approximate the electronic exchange interaction between the geminal-coupled electronic orbital and the regular electronic orbitals. The annihilation rates and densities calculated with all three RXCHF methods were very similar and agreed well with the ECG and SVM data. Thus, neglecting or approximating these exchange terms has a negligible effect on accuracy but offers substantial advantages in computational tractability. For the e^+LiH system, inclusion of the approximate exchange contribution, rather than neglecting this term, enabled the description of subtle features of the densities at essentially no additional computational cost.

As discussed in the Introduction, the NEO-XCHF and NEO-RXCHF methods are much more tractable than the ECG and SVM methods because only electron-positron correlation is treated explicitly. The results shown herein indicate that the RXCHF methods can provide accurate annihilation rates and densities, as well as other local properties. Furthermore, the RXCHF method also offers significant improvement in computational tractability over the XCHF method. Table 7.5 provides example timings for the e^+Li and LiPs systems and clearly demonstrates the advantage of RXCHF-ae and RXCHF-ne over RXCHF-fe and XCHF. Note that RXCHF-fe scales similarly to XCHF and is actually slightly slower than XCHF for the example provided in Table 7.5 because of the more complicated self-consistent-field procedure. As discussed above, however, the RXCHF-fe method is significantly more accurate than the XCHF method for positronic molecular systems, thereby justifying the slightly greater computational cost.

The RXCHF methods appear to provide a promising, computationally tractable alternative to the XCHF method. Current research focuses on extending this formalism to explicitly correlate more than a single electronic orbital to the nuclear (or positronic) orbital. An important direction will be the application of RXCHF to proton-containing systems, in which all electrons and select protons are treated quantum mechanically. In a proton transfer system, typically only a subset of electronic orbitals localized near the donor, acceptor, and transferring hydrogen would be explicitly correlated to the nuclear orbital associated with the transferring proton. In addition, electronic correlation will be included by combining RXCHF with second-order perturbation theory (MP2) or density functional theory (DFT) approaches. The present paper provides the foundation for these future directions.

b_k	γ_k
4.8792	0.015330
18.6874	0.049880
25.85601	0.16032
21.1435	0.55758
13.4063	2.3457
7.1828	13.732
2.9804	154.96
0.5494	1088.6

Table 7.1: Geminal parameters optimized using the XCHF wavefunction given by $\Psi\left(\mathbf{x}_1^e, \mathbf{x}^p\right) = \chi^e\left(\mathbf{x}_1^e\right)\chi^p\left(\mathbf{x}^p\right)g\left(\mathbf{r}_1^e, \mathbf{r}^p\right)$ for the e^-e^+X model system, where X is a fixed classical point charge of +0.5.

Method	Energy	$\langle \delta_{ep} \rangle$	Γ_a	Γ_s	Γ_t
NEO-HF	-7.4260	6.0150×10^{-6}	3.3556×10^{-4}	1.1503×10^{-3}	6.3975×10^{-5}
XCHF	-7.4756	0.032631	1.7361	6.9443	4.0981×10^{-6}
RXCHF-ne	-7.4810	0.033194	1.6759	6.6999	1.2103×10^{-3}
RXCHF-ae	-7.4818	0.032981	1.6657	6.6560	2.2227×10^{-3}
RXCHF-fe	-7.4818	0.032981	1.6656	6.6560	2.1803×10^{-3}
SVM	-7.5323	0.034698	1.7512	6.9956	3.0833×10^{-3}

Table 7.2: Quantities calculated for the positron-lithium system using an even-tempered 10s/8s basis set with GTG scheme (ii). All quantities are given in atomic units except annihilation rates given in ns^{-1}. The SVM data, as well as the prefactor of 201.8788 used to calculate annihilation rates, were obtained from Ref. 16.

Method	Energy	$\langle \delta_{ep} \rangle$	$\Gamma_{2\gamma}$
NEO-HF	-7.5260	0.0010167	0.051150
XCHF	-7.5974	0.011206	0.56559
RXCHF-ne	-7.6404	0.038572	1.9405
RXCHF-ae	-7.6428	0.038557	1.9398
SVM	-7.7386	0.04188	2.107

Table 7.3: Quantities calculated for the lithium positride system using an even-tempered 8s/6s basis set with GTG scheme (ii). All quantities are given in atomic units except annihilation rates given in ns^{-1}. The SVM data, as well as the prefactor of 100.6174809 used to calculate annihilation rates, were obtained from Ref. 19.

Method	Energy	$\langle \delta_{ep} \rangle$	$\Gamma_{2\gamma}$
NEO-HF	-7.9674	0.00061346	0.030961
XCHF	-7.9732	0.012312	0.62080
RXCHF-ne	-7.9483	0.021508	1.0844
RXCHF-ae	-7.9529	0.022100	1.1142
RXCHF-ne*	-7.9707	0.027580	1.3906
RXCHF-ae*	-7.9747	0.027616	1.3924
ECG	-8.1049	0.02499	1.26

Table 7.4: Quantities calculated for the positron-lithium hydride system with GTG scheme (ii). All quantities are given in atomic units except annihilation rates given in ns^{-1}. All values were calculated with the pc0/6s basis set except those with an asterisk, which were calculated with the pc1/6s-opt basis set. The ECG data, as well as the prefactor of 100.8403 used to calculate annihilation rates, were obtained from Ref. 12.

Method	e^+Li	LiPs
XCHF	991	405
RXCHF-fe	1150	–
RXCHF-ae	4.15	0.263
RXCHF-ne	4.08	0.255

Table 7.5: Wall timings given in minutes for complete calculations involving integral evaluation and 100 iterations of the self-consistent-field procedure. The e^+Li calculations were performed in serial with a 10s/8s basis set and $N_{\text{gem}} = 8$. The LiPs calculations were performed in parallel using eight cores via OpenMP parallelization with an 8s/6s basis set and $N_{\text{gem}} = 8$. These calculations were performed on a 3.00GHz Intel Xeon E5450 machine.

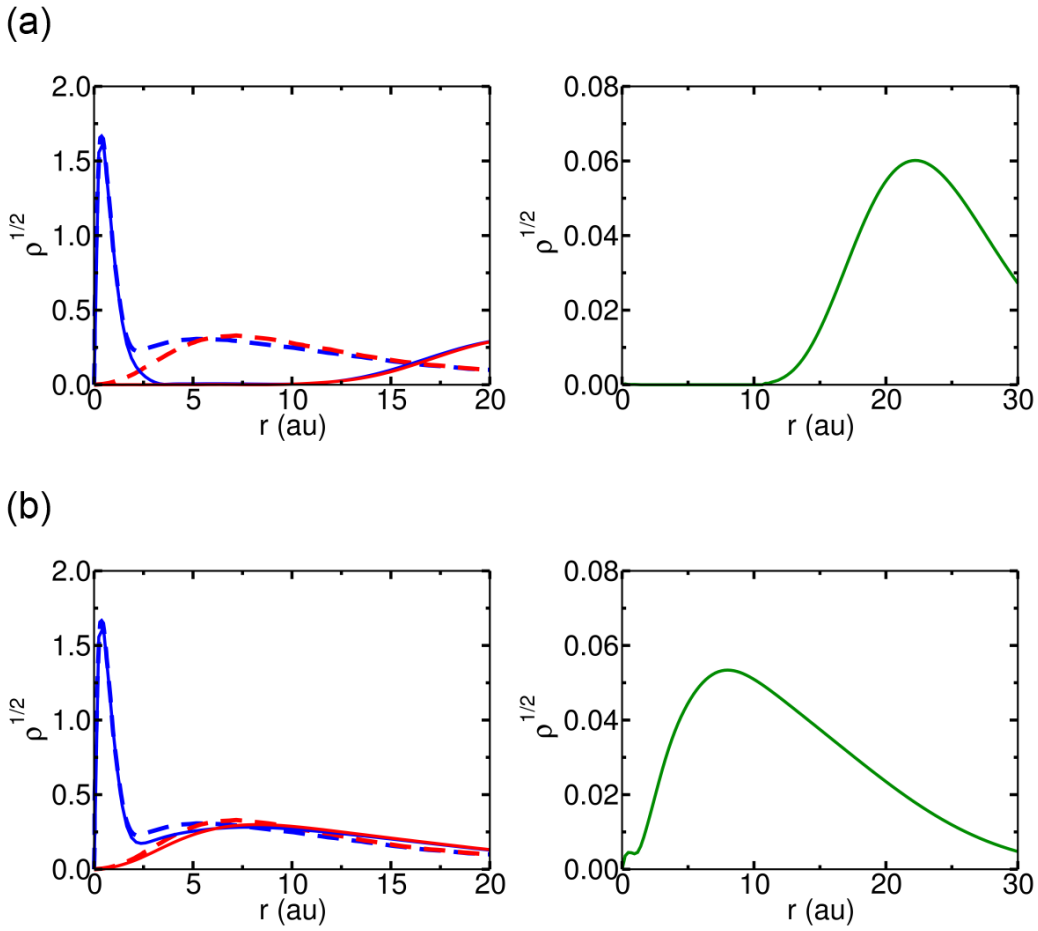

Figure 7.1: Radially-averaged electronic (blue) and positronic (red) single-particle densities and contact densities (green) calculated for the e^+Li system using (a) XCHF and (b) RXCHF-fe with GTG scheme (ii) using an even-tempered 10s/8s basis set. The dashed lines represent benchmark SVM data digitized from Ref. 19.

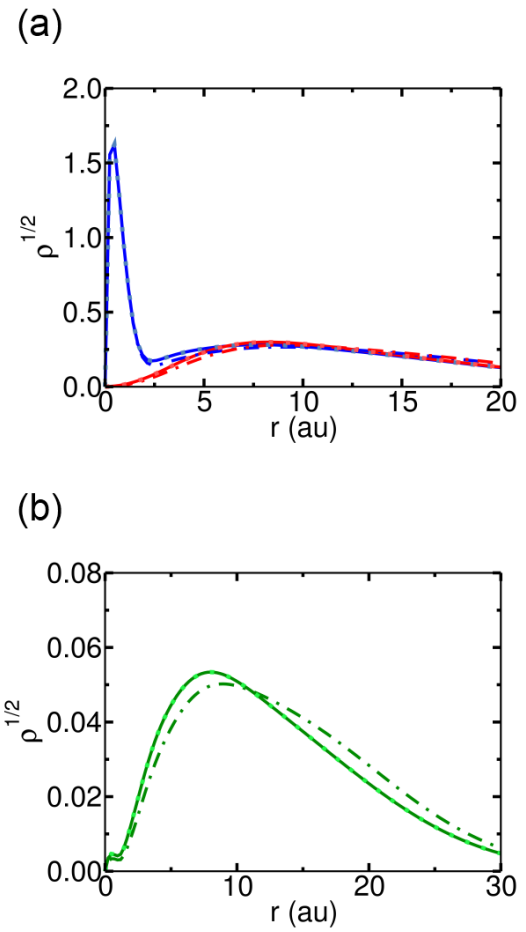

Figure 7.2: Radially-averaged (a) electronic (blue) and positronic (red) single-particle densities and (b) contact densities (green) calculated for the e^+Li system using RXCHF-fe (solid), RXCHF-ae (dotted) and RXCHF-ne (dotted-dashed) with GTG scheme (ii) using an even-tempered 10s/8s basis set. The single-particle densities from all three methods are very similar, and all densities for the RXCHF-ae and RXCHF-fe methods are virtually indistinguishable.

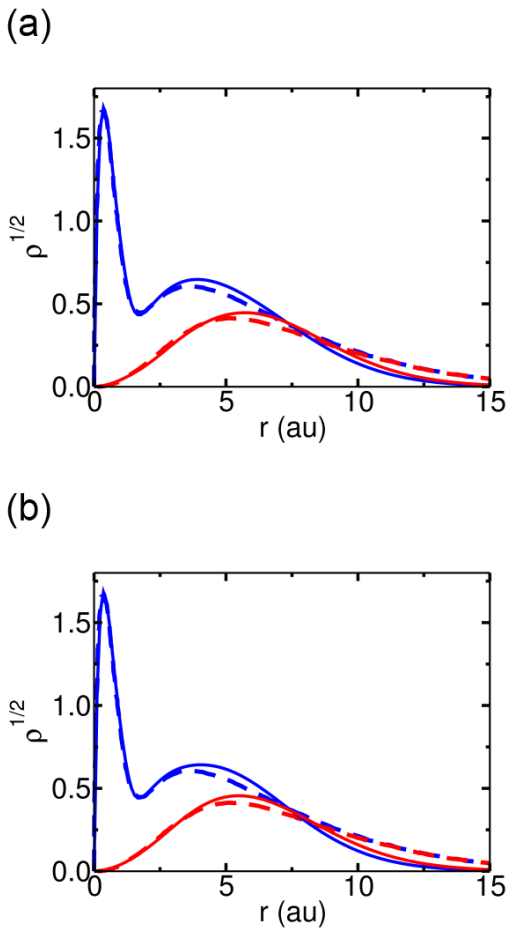

Figure 7.3: Radially-averaged electronic (blue) and positronic (red) single-particle densities calculated for the LiPs system using (a) XCHF and (b) RXCHF-ae with GTG scheme (ii) using an even-tempered 8s/6s basis set. The dashed lines represent benchmark SVM data digitized from Ref. 19.

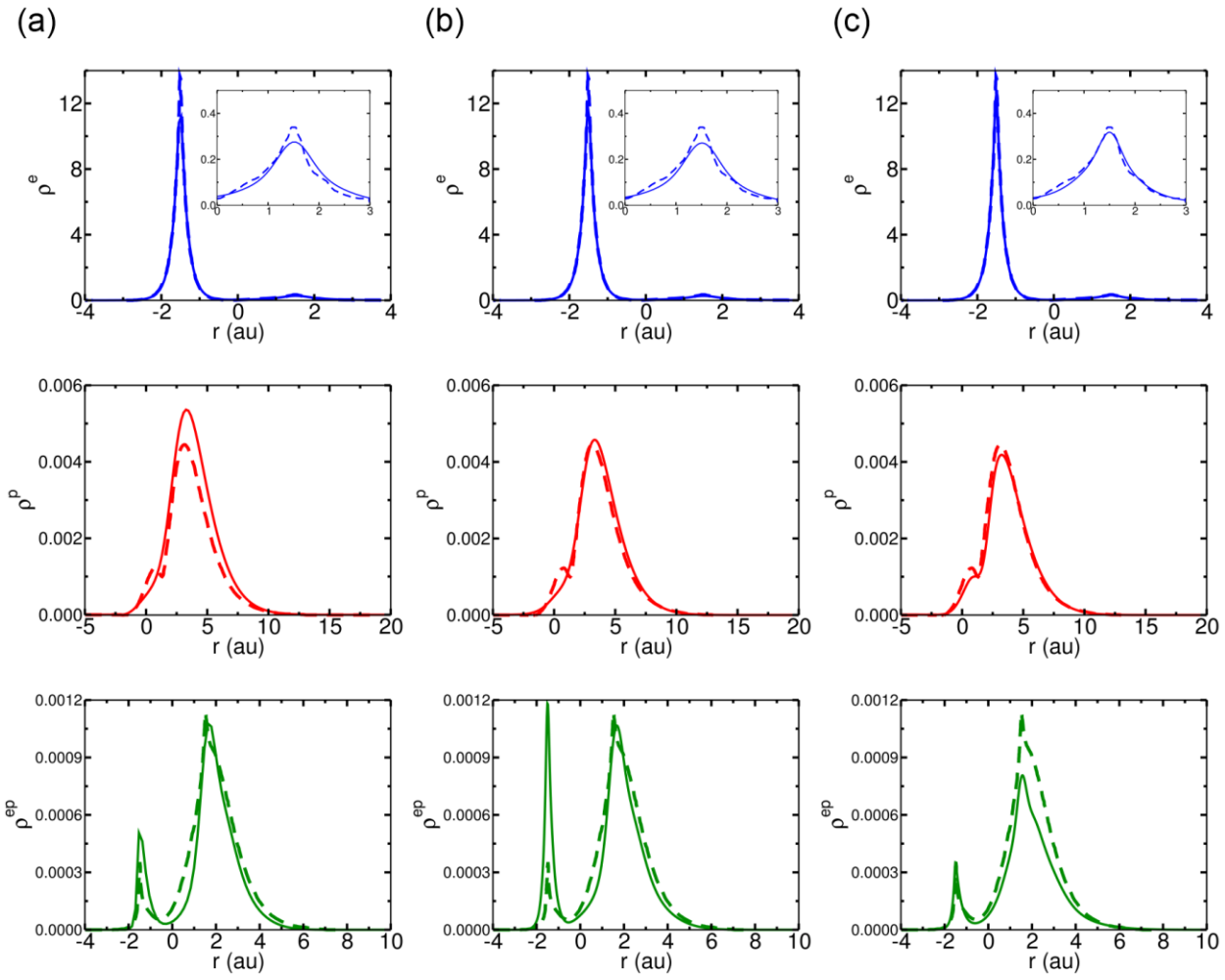

Figure 7.4: Electronic (top, blue), positronic (middle, red), and contact (bottom, green) densities calculated for the e^+LiH system using GTG scheme (ii) with various methods: (a) XCHF with the pc0/6s basis set; (b) RXCHF-ae with the pc0/6s basis set; (c) RXCHF-ae with the pc1/6s-opt basis set. The Li and H nuclei are located at $r = -1.5$ a.u. and $r = 1.5$ a.u., respectively. The dashed lines represent benchmark ECG data digitized from Ref. 12.

References

(1) R. Ruffini, G. Vereshchagin, and S. S. Xue, *Physics Reports* **487**, 1 (2010).
(2) G. F. Gribakin, J. A. Young, and C. M. Surko, *Reviews of Modern Physics* **82**, 2557 (2010).
(3) R. A. Ferrell, *Reviews of Modern Physics* **28**, 308 (1956).
(4) M. J. Puska and R. M. Nieminen, *Reviews of Modern Physics* **66**, 841 (1994).
(5) M. E. Phelps, *Proceedings of the National Academy of Sciences of the United States of America* **97**, 9226 (2000).
(6) P. E. Adamson, X. F. F. Duan, L. W. Burggraf, M. V. Pak, C. Swalina, and S. Hammes-Schiffer, *Journal of Physical Chemistry A* **112**, 1346 (2008).
(7) M. V. Pak, A. Chakraborty, and S. Hammes-Schiffer, *Journal of Physical Chemistry A* **113**, 4004 (2009).
(8) C. Swalina, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Physics* **136**, 164105 (2012).
(9) S. Bubin and L. Adamowicz, *Journal of Chemical Physics* **120**, 6051 (2004).
(10) S. Bubin and L. Adamowicz, *Physical Review A* **74**, 052502 (2006).
(11) S. Bubin, M. Pavanelo, W. C. Tung, K. L. Sharkey, and L. Adamowicz, *Chemical Reviews* **113**, 36 (2013).
(12) K. Strasburger, *Journal of Chemical Physics* **111**, 10555 (1999).
(13) K. Strasburger, *Journal of Chemical Physics* **114**, 615 (2001).
(14) K. Strasburger and H. Chojnacki, *Journal of Chemical Physics* **108**, 3218 (1998).
(15) K. Strasburger and M. Wolcyrz, *Molecular Physics* **105**, 467 (2007).
(16) J. Mitroy, *Physical Review A* **70**, 024502 (2004).
(17) J. Mitroy, *Physical Review A* **73**, 054502 (2006).
(18) J. Mitroy, M. W. J. Bromley, and G. G. Ryzhikh, *Journal of Physics B-Atomic Molecular and Optical Physics* **35**, R81 (2002).
(19) G. G. Ryzhikh, J. Mitroy, and K. Varga, *Journal of Physics B-Atomic Molecular and Optical Physics* **31**, 3965 (1998).
(20) H. Chojnacki and K. Strasburger, *Molecular Physics* **104**, 2273 (2006).
(21) N. Jiang and D. M. Schrader, *Journal of Chemical Physics* **109**, 9430 (1998).
(22) Y. Kita, R. Maezono, M. Tachikawa, M. Towler, and R. J. Needs, *Journal of Chemical Physics* **131**, 134310 (2009).
(23) H. A. Kurtz and K. D. Jordan, *Journal of Chemical Physics* **75**, 1876 (1981).
(24) I. Kylanpaa, T. T. Rantala, and D. M. Ceperley, *Physical Review A* **86**, 052506 (2012).
(25) M. Mella, G. Morosi, D. Bressanini, and S. Elli, *Journal of Chemical Physics* **113**, 6154 (2000).
(26) J. Mitroy and M. W. J. Bromley, *Physical Review A* **73** (2006).
(27) K. Strasburger, *Structural Chemistry* **15**, 415 (2004).
(28) K. Strasburger and H. Chojnacki, *Chemical Physics Letters* **241**, 485 (1995).
(29) M. Tachikawa, *Chemical Physics Letters* **350**, 269 (2001).
(30) M. Tachikawa, K. Mori, K. Suzuki, and K. Iguchi, *International Journal of Quantum Chemistry* **70**, 491 (1998).
(31) A. Takatsuka and S. Ten-no, *Bulletin of the Korean Chemical Society* **24**, 859 (2003).
(32) A. Sirjoosingh, M. V. Pak, C. Swalina, and S. Hammes-Schiffer, *Journal of Chemical Physics* **139**, 034102 (2013).

(33) F. Jensen, *Journal of Chemical Physics* **115**, 9113 (2001).
(34) F. Jensen, *Journal of Physical Chemistry A* **111**, 11198 (2007).

Chapter 8

Reduced Explicitly Correlated Hartree-Fock Approach within the Nuclear-Electronic Orbital Framework: Extension to Quantum Protons‡

8.1. Introduction

Nuclear quantum effects such as zero-point energy and hydrogen tunneling play an important role in a wide range of systems.[1-3] In addition, electron-proton nonadiabatic effects, which correspond to the breakdown of the Born-Oppenheimer separation between electrons and certain protons, can also play a significant role. In particular, these types of non-Born-Oppenheimer effects have been shown to be prevalent in proton-coupled electron transfer (PCET) reactions,[4-6] which are essential to many enzymatic reactions, molecular electrocatalysis, photosynthesis, and other energy conversion processes.[7-9] Conventional electronic structure methods that invoke the Born-Oppenheimer approximation and treat all nuclei classically are often not suitable for describing the coupling between the electrons and the transferring proton in PCET reactions. Thus, the development of theoretical methods that treat the transferring proton quantum mechanically and include non-Born-Oppenheimer effects is important for the description of these processes.

To study these types of systems, a variety of non-Born-Oppenheimer methods have been developed.[10-32] One class of approaches developed by our group is the nuclear-electronic orbital (NEO) method, in which all electrons and select nuclei are treated quantum mechanically on

equal footing using orbital-based techniques.[14,15,26-30,32] This type of *ab initio*, non-Born-Oppenheimer electronic structure approach aims to solve the time-independent Schrödinger equation for a mixed-component system. The NEO approach enables the accurate calculation of quantities that inherently include non-Born-Oppenheimer effects between the electrons and select nuclei, which are typically protons. The resulting molecular geometries, reaction paths, and couplings could be used in conjunction with existing PCET rate theories[33] to predict experimentally relevant quantities such as rate constants and kinetic isotope effects.

As discussed extensively elsewhere, mean-field approaches such as Hartree-Fock[32] and mean-field-based approaches such as MP2[30] are not adequate within the NEO framework because they do not account for the substantial electron-proton (ep) dynamical correlation.[29] Electron-proton correlation is of greater qualitative significance than electron-electron correlation because electrons and protons interact via an attractive potential and thus are often in close proximity. As a result, the nuclear densities are much too localized when calculated with mean-field-based approaches, leading to significant inaccuracies in vibrational frequencies, vibrationally averaged geometries, and geometric isotope effects.[34-36] Thus, the recent focus of our laboratory has been the development of explicitly correlated methods within the NEO framework,[14,15,26,27,29,37] inspired by many of the analogous approaches developed for electron-electron correlation.[38-52] Both explicitly correlated wavefunction methods[14,27,29,37] and density functional theory approaches with electron-proton functionals derived from these explicitly correlated wavefunctions[15,26,53] were developed within the NEO framework and have proved to be significantly more accurate than the mean-field-based approaches. However, these methods have all suffered from poor computational tractability and scaling properties, mainly due to the many-particle integrals arising from explicitly correlating the electrons and quantum nuclei.[54]

Recently, we developed the reduced explicitly correlated Hartree-Fock approach within the NEO framework, denoted NEO-RXCHF,[27] which differs from previous explicitly correlated NEO approaches, denoted NEO-XCHF,[14] in that only a single electronic orbital is explicitly correlated to the quantum nuclear orbital. We proposed such an ansatz with the principal goal of improving tractability while still maintaining the essential portion of the electron-proton dynamical correlation. We determined that significant improvement in computational tractability could be attained by neglecting or approximating parts of the electronic exchange interactions, leading to methods denoted NEO-RXCHF-ne and NEO-RXHF-ae, respectively. Specifically, we neglected or approximated the exchange interactions between the orbitals that were explicitly correlated to the nuclear orbital and those that were not explicitly correlated to the nuclear orbital.. We found that these approximations did not compromise the accuracy of the method in applications to several positron-containing molecular systems, and we showed that the NEO-RXCHF method is able to successfully describe the densities and annihilation rates for these types of systems.[55]

In this study, we extend the NEO-RXCHF approach to explicitly correlate multiple electronic orbitals to the quantum nuclear orbital. When protons are treated quantum mechanically in molecular species, at least two electronic spin orbitals should be explicitly correlated to the proton orbital, corresponding to the two electrons involved in a covalent bond (e.g., a terminal O-H bond). For a donor-hydrogen-acceptor interface, four electronic spin orbitals should be explicitly correlated to the proton orbital, corresponding to a covalent bond or hydrogen bond between the proton and the donor atom and between the proton and the acceptor atom. Therefore, we generalize the previous NEO-RXCHF methodology[27] to the case of any number of explicitly correlated orbitals. This extension becomes equivalent to the NEO-XCHF

method when all electronic orbitals are explicitly correlated to the proton orbital. We also develop approximations to certain parts of the electronic exchange terms for this more general RXCHF approach, analogous to the ideas employed in our previous work,[27] and demonstrate their effectiveness for proton-containing systems.

An outline of this paper is as follows. We first derive the RXCHF equations for a new ansatz for the mixed nuclear-electronic wavefunction that generalizes to any number of explicitly correlated electronic orbitals coupled to one quantum nuclear orbital. We derive the associated RXCHF-ne equations, which neglect electronic exchange between the two types of electronic orbitals (i.e., those that are explicitly correlated to the nuclear orbital and those that are not). We also derive the RXCHF-ae equations, which recover some of the missing exchange in the RXCHF-ne wavefunction by including the Hartree-Fock exchange between these two types of electronic orbitals. In addition to reporting all of the necessary working equations for applying these RXCHF approaches, we comment on several important points concerning the computational implementation of these methods. We then demonstrate the application of the RXCHF method to hydrogen cyanide, where all electrons and the proton are treated quantum mechanically, and only two electronic orbitals are explicitly correlated to the proton orbital. This example serves as a benchmark for applications of RXCHF to proton-containing species. In the final section, we summarize the advantages and limitations of this approach and discuss directions for future work.

8.2. Theory

We consider a system of N electrons, one quantum proton, and N_c fixed classical nuclei. The Hamiltonian in atomic units for this system is given by

$$H = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \frac{1}{2m_p}\nabla_p^2 - \sum_{i=1}^{N}\sum_{A=1}^{N_c}\frac{Z_A}{\left|\mathbf{r}_i^e - \mathbf{r}_A^c\right|} + \sum_{A=1}^{N_c}\frac{Z_A}{\left|\mathbf{r}^p - \mathbf{r}_A^c\right|} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{\left|\mathbf{r}_i^e - \mathbf{r}_j^e\right|} - \sum_{i=1}^{N}\frac{1}{\left|\mathbf{r}_i^e - \mathbf{r}^p\right|}, \tag{8.1}$$

where $\mathbf{r}^e$, $\mathbf{r}^p$, and $\mathbf{r}^c$ denote the collective spatial coordinates of the electrons, quantum proton, and classical nuclei, respectively, m_p is the mass of the proton, and Z_A is the charge of the A^{th} classical nucleus. Note that the term corresponding to repulsion between classical nuclei is omitted from Eq. (8.1) for simplicity.

In the NEO-XCHF approach,[14,29] the nuclear-electronic wavefunction is assumed to be of the form

$$\Psi^{\text{XCHF}}\left(\mathbf{x}^e, \mathbf{x}^p\right) = \Phi^e\left(\mathbf{x}^e\right)\chi^p\left(\mathbf{x}^p\right)G\left(\mathbf{r}^e, \mathbf{r}^p\right), \tag{8.2}$$

where $\Phi^e\left(\mathbf{x}^e\right) = \left|\chi_1^e\left(\mathbf{x}_1^e\right)\cdots\chi_N^e\left(\mathbf{x}_N^e\right)\right\rangle$ is a Slater determinant of N electronic spin orbitals, χ^p is a nuclear spin orbital, and

$$G\left(\mathbf{r}^e, \mathbf{r}^p\right) = \sum_{i=1}^{N} g\left(\mathbf{r}_i^e, \mathbf{r}^p\right) \tag{8.3}$$

with GTG functions defined as

$$g\left(\mathbf{r}_i^e, \mathbf{r}^p\right) = \sum_{k=1}^{N_{\text{gem}}} b_k e^{-\gamma_k\left|\mathbf{r}_i^e - \mathbf{r}^p\right|^2}. \tag{8.4}$$

In these expressions and those that follow, $\mathbf{x}^e$ and $\mathbf{x}^p$ denote the collective spin coordinates for the electrons and the quantum proton, respectively. It is immediately apparent that $\Psi^{\text{XCHF}}\left(\mathbf{x}^e, \mathbf{x}^p\right)$ is antisymmetric with respect to permutations among the electronic indices. Note that the ansatz in Eq. (8.2) differs from that in Ref. 14, which has a factor of $1+G$ rather than G, but we proceed with this form because the resulting expressions presented below are

significantly simplified. To recover the $1+G$ form of the wavefunction, an additional geminal can be included to mimic the term of unity by choosing germinal parameters $b_k = 1/N$ and $\gamma_k \approx 0$.

In the NEO-RXCHF approach, we partition the N electron, 1 quantum proton system into two subsystems: an N_r electron system and an N_s electron, 1 quantum proton system ($N_r + N_s = N$). We then invoke the following ansatz for the nuclear-electronic wavefunction:

$$\begin{aligned}\Psi^{\text{RXCHF-ne}}\left(\mathbf{x}_1^e,\ldots\mathbf{x}_N^e,\mathbf{x}^p\right) &= \Psi^{\text{RXCHF-ne}}\left(\mathbf{x}_1^r,\ldots\mathbf{x}_{N_r}^r,\mathbf{x}_1^s,\ldots\mathbf{x}_{N_s}^s,\mathbf{x}^p\right)\\ &= \Phi^{\text{e}}\left(\mathbf{x}_1^r,\ldots\mathbf{x}_{N_r}^r\right)\Psi^{\text{XCHF}}\left(\mathbf{x}_1^s,\ldots\mathbf{x}_{N_s}^s,\mathbf{x}^p\right)\\ &= \mathcal{A}\left\{\chi_1^r\left(\mathbf{x}_1^r\right),\ldots,\chi_{N_r}^r\left(\mathbf{x}_{N_r}^r\right)\right\}\\ &\quad\times\mathcal{A}\left\{\chi_1^s\left(\mathbf{x}_1^s\right),\ldots,\chi_{N_s}^s\left(\mathbf{x}_{N_s}^s\right)\right\}\chi^p\left(\mathbf{x}^p\right)G\left(\mathbf{r}_1^s,\ldots\mathbf{r}_{N_s}^s,\mathbf{r}^p\right),\end{aligned} \tag{8.5}$$

where the N_s electronic spin orbitals that comprise the XCHF part of the wavefunction, $\left\{\chi_i^s\right\}$, are geminal-coupled to the nuclear spin orbital, χ^p, while the N_r electronic spin orbitals that comprise the Hartree-Fock part of the wavefunction, $\left\{\chi_i^r\right\}$, are not geminal-coupled to the nuclear spin orbital. In this expression, $\mathcal{A}$ represents the antisymmetrization operator. Applying this ansatz physically distinguishes between two subsets of electrons: the N_s electrons that are explicitly correlated to the quantum proton are denoted "special" electrons, and the N_r electrons that are not explicitly correlated are denoted "regular" electrons. Antisymmetry is maintained within each subsystem (e.g., exchange among the special electrons and among the regular electrons is fully included), but the special electrons are distinguished from the regular electrons and are effectively different types of particles. We denote this ansatz as RXCHF-ne to emphasize

that this approach neglects exchange between the special and regular electrons. The ansatz in Eq. (8.5) was developed in Ref. 27 for the specific case of $N_s = 1$.

The energy associated with the RXCHF wavefunction is defined as an expectation value over the Hamiltonian operator given in Eq. (8.1),

$$\begin{aligned} E^{\text{RXCHF-ne}} &= \frac{\left\langle \Psi^{\text{RXCHF-ne}} \middle| H \middle| \Psi^{\text{RXCHF-ne}} \right\rangle}{\left\langle \Psi^{\text{RXCHF-ne}} \middle| \Psi^{\text{RXCHF-ne}} \right\rangle} \\ &= E^{\text{HF}} + E^{\text{XCHF}} + E^{\text{int-ne}}, \end{aligned} \tag{8.6}$$

where the partition of the RXCHF-ne energy follows from the ansatz in Eq. (8.5). In Eq. (8.6), E^{HF} is the Hartree-Fock energy of an N_r-electron system, E^{XCHF} is the XCHF energy of an N_s-electron, 1 quantum proton system, and $E^{\text{int-ne}}$ is the interaction energy between these two subsystems. These quantities are explicitly defined in subsequent sections.

Analogous to the previous RXCHF formalism,[27] we may formulate an approach that accounts for some of the missing exchange terms between regular and special electrons. To this end, we define an RXCHF-ae energy, denoting approximate exchange, as

$$E^{\text{RXCHF-ae}} = E^{\text{RXCHF-ne}} + E^{\text{ex}} \equiv \text{E}^{\text{HF}} + \text{E}^{\text{XCHF}} + \text{E}^{\text{int-ae}}, \tag{8.7}$$

where the RXCHF-ne energy of Eq. (8.6) is corrected by an exchange term, E^{ex}, constructed to preserve the same scaling as the RXCHF-ne approach. This correction is included in the interaction term as $E^{\text{int-ae}} = E^{\text{int-ne}} + E^{\text{ex}}$ and includes the Hartree-Fock level of exchange between regular and special electron orbitals. Note that the addition of E^{ex} to the energy implies that the RXCHF-ae energy does not arise from a specific wavefunction; thus, approximations need to be invoked to evaluate expectation values of other operators.

Varying Eq. (8.6) or (8.7) with respect to the spin orbitals leads to modified Hartree-Fock (HF) equations of the form

$$f^r \chi_i^r = \varepsilon_i^r \chi_i^r + \sum_{j=1}^{N_s} \varepsilon_{ij}^r \chi_j^s, \quad i \in \{1,\ldots,N_r\} \tag{8.8}$$

$$f^s \chi_i^s = \varepsilon_i^s \chi_i^s + \sum_{j=1}^{N_r} \varepsilon_{ij}^s \chi_j^r, \quad i \in \{1,\ldots,N_s\} \tag{8.9}$$

$$f^p \chi^p = \varepsilon^p \chi^p \tag{8.10}$$

for the regular electrons, the special electrons, and the quantum proton, respectively. Since variations of the energy with respect to regular, special, or quantum proton spin orbitals are distinct from each other, three coupled Fock equations appear instead of one Fock equation as in standard Hartree-Fock theory.[56] The presence of off-diagonal Lagrange multipliers, $\{\varepsilon_{ij}^r\}$ and $\{\varepsilon_{ij}^s\}$, in Eqs. (8.8) and (8.9) constrain the regular and special electron orbitals to remain orthogonal to each other. The terms involving these multipliers can be rigorously eliminated while still preserving orthonormality, as discussed in Section 8.3.

The evaluation of E^{HF} is found in standard quantum chemistry textbooks,[56] and the derivation of E^{XCHF} is discussed in detail in Ref. 14. As discussed above, the XCHF ansatz as defined in Eq. (8.2) differs from that developed previously in that the correlating factor is G rather than $1+G$. However, the equations presented in Ref. 14 can be directly utilized by removing all terms involving operators with zero or one geminal factor, g. For convenience, we report the definitions of E^{HF} and E^{XCHF} in the Section G.4. In the following subsections, we describe the evaluation of the interaction energy term E^{int} in Eq. (8.6) and the evaluation of the approximate exchange term E^{ex} in Eq. (8.7) for the RXCHF methods. We also present a restricted Hartree-Fock (RHF) formalism, in which all electrons of opposite spin are paired, and report working expressions for the RXCHF methods.

A. Overlap

The overlap of the RXCHF-ne wavefunction is evaluated as $S^{\text{RXCHF-ne}} \equiv \left\langle \Psi^{\text{RXCHF-ne}} \middle| \Psi^{\text{RXCHF-ne}} \right\rangle = \left\langle \Psi^{\text{XCHF}} \middle| \Psi^{\text{XCHF}} \right\rangle \left\langle \Phi^{e} \middle| \Phi^{e} \right\rangle = \left\langle \Psi^{\text{XCHF}} \middle| \Psi^{\text{XCHF}} \right\rangle$, which, taken from Ref. 14 and using the modifications described above, is given by

$$\begin{aligned} S^{\text{RXCHF-ne}} &= \sum_{a=1}^{N_s} \left\langle \chi^p(p)\chi_a^s(1) \middle| g(1,p)^2 \middle| \chi^p(p)\chi_a^s(1) \right\rangle \\ &\quad + \sum_{a=1}^{N_s}\sum_{b=1}^{N_s} \Big(\left\langle \chi^p(p)\chi_a^s(1)\chi_b^s(2) \middle| g(1,p)\,g(2,p) \middle| \chi^p(p)\chi_a^s(1)\chi_b^s(2) \right\rangle \\ &\quad - \left\langle \chi^p(p)\chi_a^s(1)\chi_b^s(2) \middle| g(1,p)\,g(2,p) \middle| \chi^p(p)\chi_b^s(1)\chi_a^s(2) \right\rangle \Big) \\ &\equiv \sum_{a=1}^{N_s} \left\langle \chi^p\chi_a^s \middle| g^2 \middle| \chi^p\chi_a^s \right\rangle \\ &\quad + \sum_{a=1}^{N_s}\sum_{b=1}^{N_s} \Big(\left\langle \chi^p\chi_a^s\chi_b^s \middle| g(1,p)\,g(2,p) \middle| \chi^p\chi_a^s\chi_b^s \right\rangle - \left\langle \chi^p\chi_a^s\chi_b^s \middle| g(1,p)\,g(2,p) \middle| \chi^p\chi_b^s\chi_a^s \right\rangle \Big). \end{aligned} \tag{8.11}$$

In these expressions and those that follow, angular brackets denote integration over all coordinates. Spin-coordinate dependence of the orbitals is abbreviated as p for $\mathbf{x}^p$ and i for $\mathbf{x}_i^r$ or $\mathbf{x}_i^s$, and the distinction between regular and special coordinates will be inferred by the orbitals depending on these coordinates. The spatial coordinates appearing in the geminal functions, which depend on only the special electron coordinates, will also be abbreviated as i for $\mathbf{r}_i^s$. In addition, we write $g \equiv g(1,p)$, and the coordinate dependence of the spin orbitals appearing in bras or kets follows the order $p,1,2,\ldots$, e.g., $\left|\chi^p\chi_a^s\chi_b^s\right\rangle \equiv \left|\chi^p(p)\chi_a^s(1)\chi_b^s(2)\right\rangle$.

B. Energy

The interaction energy associated with the RXCHF-ne wavefunction can be expressed as

$$E^{\text{int-ne}} = \frac{E_2 + E_3 + E_4}{S^{\text{RXCHF-ne}}}, \tag{8.12}$$

where the overlap is given by Eq. (8.11).

The second-order terms are given by

$$E_2 = \sum_{a=1}^{N_r}\sum_{b=1}^{N_s}\left\langle \chi^p \chi_a^r \chi_b^s \left| \Omega_2\left(p,1,2\right)\right| \chi^p \chi_a^r \chi_b^s \right\rangle, \tag{8.13}$$

where

$$\Omega_2\left(p,1,2\right) = g\left(2,p\right)\left[V_{ee}\left(1,2\right)+V_{ep}\left(p,1\right)\right]g\left(2,p\right). \tag{8.14}$$

In this expression and those that follow, $V_{ee}\left(i,j\right) = \left|\mathbf{r}_i - \mathbf{r}_j\right|^{-1}$ is the Coulomb repulsion operator between the i^{th} and j^{th} electrons, and $V_{ep}\left(p,i\right) = -\left|\mathbf{r}^p - \mathbf{r}_i\right|^{-1}$ is the Coulomb attraction operator between the quantum nucleus and the i^{th} electron. The electron coordinates correspond to regular or special electrons inferred by the coordinate dependence of the corresponding orbitals.

The third-order terms are given by

$$\begin{aligned} E_3 = \sum_{a=1}^{N_r}\sum_{b=1}^{N_s}\sum_{c=1}^{N_s}\Big(&\left\langle \chi^p \chi_a^r \chi_b^s \chi_c^s \left| \Omega_3\left(p,1,2,3\right)\right| \chi^p \chi_a^r \chi_b^s \chi_c^s \right\rangle \\ &-\left\langle \chi^p \chi_a^r \chi_b^s \chi_c^s \left| \Omega_3\left(p,1,2,3\right)\right| \chi^p \chi_a^r \chi_c^s \chi_b^s \right\rangle\Big), \end{aligned} \tag{8.15}$$

where

$$\Omega_3\left(p,1,2,3\right) = g\left(2,p\right)\left[V_{ee}\left(1,3\right)\right]g\left(2,p\right) + g\left(2,p\right)\left[2V_{ee}\left(1,2\right)+V_{ep}\left(p,1\right)\right]g\left(3,p\right). \tag{8.16}$$

The fourth-order terms are given by

$$E_4 = \sum_{a=1}^{N_r}\sum_{b=1}^{N_s}\sum_{c=1}^{N_s}\sum_{d=1}^{N_s}\left\langle \chi^p \chi_a^r \chi_b^s \chi_c^s \chi_d^s \left| \Omega_4\left(p,1,2,3,4\right)\sum_{m=1}^{3!}\left(-1\right)^{p_m} P_m^{(3)}\left(2,3,4\right)\right| \chi^p \chi_a^r \chi_b^s \chi_c^s \chi_d^s \right\rangle, \tag{8.17}$$

where $P_m^{(3)}\left(i,j,k\right)$ is the m^{th} third-order permutation with parity p_m acting on coordinates $\left(i,j,k\right)$, and

$$\Omega_4\left(p,1,2,3,4\right) = g\left(2,p\right)\left[V_{ee}\left(1,4\right)\right]g\left(3,p\right). \tag{8.18}$$

To maintain the RXCHF-ne level of tractability while accounting for exchange interactions between special and regular electrons, we define the exchange correction for the RXCHF-ae method as

$$
\begin{aligned}
E^{\mathrm{ex}} = -\frac{1}{S^{\mathrm{RXCHF\text{-}ne}}}\Bigg(& \sum_{a=1}^{N_r}\sum_{b=1}^{N_s}\left\langle \chi^p \chi_a^r \chi_b^s \left| \Omega_2^{(\mathrm{ex})}(p,1,2) \right| \chi^p \chi_b^s \chi_a^r \right\rangle \\
& + \sum_{a=1}^{N_r}\sum_{b=1}^{N_s}\sum_{c=1}^{N_s} \\
& \Big[\left\langle \chi^p \chi_a^r \chi_b^s \chi_c^s \left| \Omega_3^{(\mathrm{ex},1)}(p,1,2,3) \right| \chi^p \chi_b^s \chi_a^r \chi_c^s \right\rangle - \left\langle \chi^p \chi_a^r \chi_b^s \chi_c^s \left| \Omega_3^{(\mathrm{ex},1)}(p,1,2,3) \right| \chi^p \chi_c^s \chi_a^r \chi_b^s \right\rangle \\
& + \left\langle \chi^p \chi_a^r \chi_b^s \chi_c^s \left| \Omega_3^{(\mathrm{ex},2)}(p,1,2,3) \right| \chi^p \chi_c^s \chi_b^s \chi_a^r \right\rangle - \left\langle \chi^p \chi_a^r \chi_b^s \chi_c^s \left| \Omega_3^{(\mathrm{ex},2)}(p,1,2,3) \right| \chi^p \chi_b^s \chi_c^s \chi_a^r \right\rangle \Big] \Bigg),
\end{aligned}
\tag{8.19}
$$

where

$$\Omega_2^{(\mathrm{ex})}(p,1,2) = g(2,p)\left[V_{ee}(1,2)\right]g(1,p), \tag{8.20}$$

$$\Omega_3^{(\mathrm{ex},1)}(p,1,2,3) = 2g(2,p)\left[V_{ee}(1,2)\right]g(3,p), \tag{8.21}$$

$$\Omega_3^{(\mathrm{ex},2)}(p,1,2,3) = g(2,p)\left[V_{ee}(1,3)\right]g(2,p). \tag{8.22}$$

This definition of E^{ex} includes Hartree-Fock exchange, provided that the $1+G$ ansatz is invoked by appropriately choosing geminal parameters. The inclusion of four-particle terms does not exceed the scaling of RXCHF-ne but ensures that the NEO-HF limit can be recovered when the geminal factor *G* approaches unity. As described above, we include this exchange term in the definition of the RXCHF-ae interaction energy, and applying Eq. (8.12), we obtain

$$E^{\mathrm{int-ae}} = \frac{E_2 + E_3 + E_4}{S^{\mathrm{RXCHF\text{-}ne}}} + E^{\mathrm{ex}}. \tag{8.23}$$

C. Fock operators

In the RXCHF-ne or RXCHF-ae approaches, the three Fock operators in Eqs. (8.8) to (8.10) are formed by varying the total energy in Eq. (8.6) or Eq. (8.7), respectively, with respect

to the regular, special, or quantum proton spin orbitals. Based on the partition of energy in these equations, and exploiting the fact that E^{HF} is independent of the special electron orbitals and proton orbital, while E^{XCHF} is independent of the regular electron orbitals, the resulting Fock operators can be expressed as

$$f^{r} = f^{r,\mathrm{HF}} + f^{r,\mathrm{int}}, \tag{8.24}$$

$$f^{s} = f^{s,\mathrm{XCHF}} + f^{s,\mathrm{int}}, \tag{8.25}$$

$$f^{p} = f^{p,\mathrm{XCHF}} + f^{p,\mathrm{int}}. \tag{8.26}$$

Here, $\delta^{r} E^{\mathrm{HF}} \to f^{r,\mathrm{HF}}$, $\delta^{s} E^{\mathrm{XCHF}} \to f^{s,\mathrm{XCHF}}$, and $\delta^{p} E^{\mathrm{int}} \to f^{p,\mathrm{int}}$, where δ^{r} denotes variations with respect to the regular spin orbitals and δ^{s} and δ^{p} are defined analogously. The Fock operators for the HF or XCHF subsystems are obtained from previously derived results[14,56] with analogous changes for the XCHF expressions as discussed above. In the following subsection, we present the contributions to the Fock operators arising from the interaction energies $E^{\mathrm{int-ne}}$ and $E^{\mathrm{int}-a\mathrm{e}}$. The resulting expressions are given in Section G.2.1.

D. Spatial Fock operators

The energies and Fock operators in the previous subsection are written in terms of the regular electron, special electron, and proton spin orbitals. In this subsection, we present expressions for the restricted closed-shell case, analogous to the restricted Hartree-Fock (RHF) approach, where we define spatial orbitals for the regular and special electrons and assume that they are doubly occupied, each with an alpha and beta spin. For example, we write each regular electron spin orbital as

$$\chi_i^r\left(\mathbf{x}_1^r\right)=\chi_i^r\left(\mathbf{r}_1^r,\omega_1^r\right)=\begin{cases}\psi_j^r\left(\mathbf{r}_1^r\right)\alpha\left(\omega_1^r\right) & i \text{ odd}\\ \psi_j^r\left(\mathbf{r}_1^r\right)\beta\left(\omega_1^r\right) & i \text{ even}\end{cases}, \tag{8.27}$$

where $i\in\left\{1,\ldots,N^r\right\}$ and $j\in\left\{1,\ldots,N^r/2\right\}$. Analogous expressions are used for the special electron spin orbitals.

Equipped with Eq. (8.27), the spatial Fock operators and their corresponding eigenvalue equations can be derived by integrating Eqs. (8.8) to (8.10) over the spins of the electrons, as well as the spin of the quantum nucleus. For the subsequent formulations, the spatial part of the quantum nuclear spin orbital is denoted ψ^p without specifying the spin of this particle because integration over the spin of this orbital always leads to unity in expectation value expressions for spin-independent operators. Again, the HF and XCHF parts of the energy and Fock operators can be obtained directly from previous literature,[14,56] and we report the closed-shell analogues of $f^{r,\text{int}}$, $f^{s,\text{int}}$ and $f^{p,\text{int}}$ in Section G.2.2. Note that this closed-shell formalism is inapplicable to open-shell systems with an odd number of regular or special electrons. Analogs of the unrestricted Hartree-Fock (UHF) and restricted open-shell Hartree-Fock (ROHF) approaches will be derived in subsequent work to consider these cases.

E. Modified Hartree-Fock-Roothaan expressions

To generate working expressions, we define atomic orbital (AO) basis sets for the electrons and the quantum proton and derive the analogs of Hartree-Fock-Roothaan equations for RXCHF.

We expand the electronic spatial orbitals in a common set of N_{ebf} electronic basis functions. The regular electron orbitals are expanded as

$$\psi_i^r\left(\mathbf{r}_1^r\right)=\sum_{\mu=1}^{N_{\text{ebf}}} C_{\mu,i}^r \phi_\mu^e\left(\mathbf{r}_1^r\right),\ i\in\{1,\ldots,N_r/2\}\ , \tag{8.28}$$

while the special electron orbitals are expanded as

$$\psi_i^s\left(\mathbf{r}_1^s\right)=\sum_{\mu"=1}^{N_{\text{ebf}}} C_{\mu",i}^s \phi_{\mu"}^e\left(\mathbf{r}_1^s\right),\ i\in\{1,\ldots,N_s/2\}\ . \tag{8.29}$$

Similarly, we expand the nuclear spatial orbital in a set of N_{pbf} nuclear basis functions as

$$\psi^p\left(\mathbf{r}^p\right)=\sum_{\mu'=1}^{N_{\text{pbf}}} C_{\mu'}^p \phi_{\mu'}^p\left(\mathbf{r}^p\right). \tag{8.30}$$

We are currently exploring the possibility of restricting the AO basis set of the special electron orbitals to a subset of the total AO basis set that is used for the regular electrons (i.e., restricting the summation in Eq. (8.29)) to gain significant advantages in computational tractability. Further discussion regarding this approach is provided in Section 8.4.

Expanding the spatial Fock operators discussed in the previous subsection in terms of the AO basis sets defined above provides working expressions of the form $\mathbf{F}^r\mathbf{C}^r=\mathbf{S}^e\mathbf{C}^r\mathbf{E}^r$, $\mathbf{F}^s\mathbf{C}^s=\mathbf{S}^e\mathbf{C}^s\mathbf{E}^s$, and $\mathbf{F}^p\mathbf{C}^p=\mathbf{S}^p\mathbf{C}^p\mathbf{E}^p$ for the regular electron, special electron, and quantum proton equations, respectively. The Fock matrices in the AO bases, $\mathbf{F}$, the MO coefficient matrices, $\mathbf{C}$, and the eigenvalue matrices, $\mathbf{E}$, are defined for each type of orbital, while the overlap matrices $\mathbf{S}^e$ and $\mathbf{S}^p$ are defined in terms of the AO basis sets given in Eqs. (8.28) and (8.30), respectively. The expressions for the Fock matrices are given in Section G.2.3.

8.3. Computational Considerations

In this section, we address several technical aspects of our implementation of the RXCHF methods described in the previous section. We discuss the scheme used to maintain orthogonality

between regular and special electron orbitals, as well as the impact of this scheme on the choice of converger for the self-consistent-field (SCF) procedure. In addition, we describe the evaluation of the many-particle integrals outlined in Appendix G, and as these constitute the bottleneck of the method, we briefly comment on the parallelization protocols of our current in-house code.

A. Orthogonality

In the RXCHF methods, the three coupled Fock equations in Eqs. (8.8) to (8.10) are solved self-consistently. As discussed previously, the expressions for the regular and special electron orbitals (Eqs. (8.8) and (8.9)) include terms involving off-diagonal Lagrange multipliers, $\left\{\varepsilon_{ij}^{r}\right\}$ and $\left\{\varepsilon_{ij}^{s}\right\}$, to preserve orthogonality between these two subsystems. The appearance of these terms prevents the application of the canonical approach to solve each equation, that is, the use of the pseudo-eigenvalue equations of the form $f\chi = \varepsilon\chi$ for each subsystem.

One method that has been developed to address this issue is the orthogonality constrained basis set expansion (OCBSE) method,[57] which was developed to address situations in which orthogonality between two subsystems, such as two generalized valence bond (GVB) shells, must be imposed. We provided a detailed discussion of our application of this method in our previous study,[27] and herein briefly comment on the differences of our implementation for this work.

Previously, we considered the case of only one special electron orbital ($N_s = 1$) and chose to impose no orthogonality conditions between regular and special electron orbitals in the Fock procedure for the regular orbitals (i.e., we did not include the terms including the $\left\{\varepsilon_{ij}^{r}\right\}$ in

Eq. (8.8)). Instead, we imposed this orthogonality condition on only the special orbital and included all terms in Eq. (8.9). Application of one part of the OCBSE procedure as described in Ref. 58 involved projecting Eq. (8.9) onto the space spanned by the regular virtual orbitals, thereby rigorously eliminating all terms involving the $\left\{\varepsilon_{ij}^{s}\right\}$. Thus, the variational procedure for the regular electron orbitals was completely unrestricted, and at each iteration of the SCF procedure, the special electron orbital was constrained to be orthogonal to the regular electron orbitals of that iteration. This procedure is effective in cases where the special electron molecular orbital (MO) is similar to a virtual regular electron MO. However, it is problematic in cases where the special MO is similar to an occupied regular MO because the special would be artificially constrained to lie in the orthogonal complement of the physically meaningful solution space.

To circumvent these issues, we revert to the original formulation of the OCBSE procedure, where the terms involving the $\left\{\varepsilon_{ij}^{r}\right\}$ are still included in Eq. (8.8). At each iteration, this equation is projected onto the space spanned by the special virtual electron orbitals of the previous iteration, thus rigorously eliminating these terms. The analogous projection of Eq. (8.9) is subsequently carried out, guaranteeing that at each iteration, the regular and special electron orbitals are orthogonal to each other. Applying this approach restricts the variational procedure for the regular electrons and prevents the occurrence of the problematic situation described above.

Finally, we note that this implementation of OCBSE does not include mixing between the occupied regular and special electron orbitals of a given iteration. This mixing can be included by varying these orbitals simultaneously using, for example, pairwise rotations involving a combined Fock operator for both subsystems. Such mixing has been discussed extensively with

respect to GVB approaches,[57,59,60] and its inclusion is expected to improve convergence. The application of such methods to RXCHF is currently under development.

B. SCF convergence

The method for solving the analogous Hartree-Fock-Roothaan equations for RXCHF is as follows. For a given quantum proton density, which is represented by MO coefficients, we solve the Fock equations for the electronic subsystem. Given regular and special electron densities from iteration $i-1$, the Fock operators for each are formed in the electronic AO basis using the expressions in Appendix G. The regular electronic Fock operator is transformed to an appropriate basis, denoted $W_r^{(i)}$, using the OCBSE procedure as described above, and is diagonalized, generating an updated regular electron density that corresponds to an orthonormal set of regular MOs at iteration i orthogonal to the occupied special MOs at iteration $i-1$. Subsequently, the special electronic Fock operator is transformed to an appropriate basis, denoted $W_s^{(i)}$, using the OCBSE procedure, and is diagonalized, generating an updated special electron density that corresponds to an orthonormal set of special MOs at iteration i orthogonal to the occupied regular MOs at iteration i. Repeated iterations over the regular and special electronic subsystem results in converged regular and special electronic densities corresponding to the fixed quantum proton density. The latter is then updated through a diagonalization step, and the procedure is repeated for the electronic subsystem. Convergence is reached when all density changes fall below a specified threshold.

This "straight diagonalization" approach mirrors that of the standard method to solve the Hartree-Fock-Roothaan equations in regular HF theory. However, we found that such an approach is not effective for the electronic subsystem, as we obtained very poor convergence

(i.e., oscillatory behavior) even for simple model systems. Thus, we applied a variant of an approximate second-order convergence scheme that has been previously developed for RHF, GVB, and MCSCF wavefunctions.[61] In this approach, MOs are rotated by an exponential transformation generating independent parameters corresponding to rotations between occupied and virtual MOs.[62] The variational optimization of these parameters can be performed by applying the Newton-Raphson procedure using an approximate inverse Hessian,[61,63]

$$\delta_i = -\mathbf{H}_i \nabla_i, \tag{8.31}$$

where the equation is written at iteration i for the displacement vector, δ_i, the approximate inverse Hessian, $\mathbf{H}_i$, and the gradient, ∇_i, all defined in terms of the independent rotation parameters obtained after the exponential transformation described above.

At each iteration, the inverse Hessian, $\mathbf{H}_i$, is approximated by augmenting the inverse Hessian of the previous iteration, $\mathbf{H}_{i-1}$, with additional information from quantities calculated in the previous iterations using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm.[64] The implementation described in Refs. 61 and 63 utilizes a strategy to apply such a procedure without storing the full Hessian matrix at any iteration; instead a few smaller quantities are stored for *all* previous iterations, thus greatly reducing storage requirements.

Implementing this type of converger for RXCHF requires application of this procedure to both the regular and special electronic variational problems separately. However, the OCBSE steps for each subsystem involve transforming the respective Fock equations to new bases that change at each iteration. To apply the second-order method described above, we must first transform any quantities required from previous iterations into the new bases. In general, the bases for different iterations do not span the same space, so we must project any quantities from previous iterations onto the bases of the current iteration. Any loss of information caused by this

projection impacts $\mathbf{H}_i$, and thus affects only the magnitude of the Newton-Raphson step in Eq. (8.31), whereas the direction of the step, ∇_i, is calculated exactly. Because we are using an approximate inverse Hessian already, we do not expect any additional difficulties arising from this projection procedure. However, the procedure of Refs. 61 and 63, which involves storing quantities from all previous iterations, becomes problematic because it would require the quantities from all previous iterations to be transformed to the new bases and, as an example, we would expect that $W_r^{(1)}$ would be significantly different than $W_r^{(i)}$ for $i \gg 1$, at least more so than would be $W_r^{(i-1)}$ from $W_r^{(i)}$. Therefore, rather than using this procedure,[61] we directly apply the BFGS algorithm and store the full Hessian matrices, $\mathbf{H}_i$ and $\mathbf{H}_{i-1}$, at each iteration i. The increased storage requirements are not an issue for the smaller systems we are currently studying and are significantly less than the storage requirements of the many-particle integrals discussed below. For the current systems of interest, a method that requires only the transformation between OCBSE bases of adjacent iterations is more accurate than a method that requires transformations among the bases of all previous iterations.

The construction of an initial guess for the MOs in RXCHF calculations is another issue that is relevant to convergence. In principle, any initial guess could be used as long as the regular and special electron MOs are orthogonal. However, we found difficulty in convergence for some of the larger model systems when standard guess orbitals, such as those obtained from diagonalizing the core Fock operator, were used. Instead we utilize a different approach, wherein the guess regular and special electron MOs are obtained from a set of localized MOs. These localized MOs can be calculated from converged Hartree-Fock orbitals using standard methods such as Foster-Boys[65] or Edmiston-Ruedenberg[66] localization.. Examination of these orbitals allows the identification of a subset that would be suitable as the initial guess for the special

electron MOs. For proton-containing system, the electronic MOs corresponding to covalent bonds or hydrogen-bonding interactions involving the quantum hydrogen nucleus are chosen to be the special electron MOs. The remaining localized orbitals can be used as the initial guess for the regular electron MOs. We have observed improved convergence when localized orbitals are used as the initial guess for the MOs. An example of this strategy is discussed in Section 8.4 for HCN.

C. Integral evaluation and parallelization

The evaluation of the RXCHF-ne and RXCHF-ae energy and corresponding Fock operators requires the calculation and storage of three-, four-, and five-particle integrals involving geminal functions, in addition to the standard two-electron integrals required for HF calculations. These integrals are listed in Appendix G. As described previously, we currently evaluate these integrals using an extension of the McMurchie-Davidson approach[67] to many-particle geminal integrals from the work of Persson and Taylor.[68] Recently, Komornicki and King developed a general, heuristic framework that enables the formulation of expressions evaluating many-particle geminal integrals over many types of operators for any number of particles.[69] Application of their method, based on a Rys polynomial approach,[70] to the evaluation of the RXCHF integrals is expected to provide further improvements in tractability, as are developments using resolution of the identity (RI) approximations. These improvements are directions of current research.

The evaluation of the many-particle integrals constitutes the bottleneck of the RXCHF approach because of the sheer number of integrals required for the evaluation of matrix elements. When $N_s \leq 2$, the bottleneck is the evaluation of the $3N_{\mathrm{pbf}}^2 N_{\mathrm{ebf}}^6$ four-particle integrals, and when

$N_s \geq 4$, the bottleneck is the evaluation of the $2N_{\text{pbf}}^2 N_{\text{ebf}}^8$ five-particle integrals, which is equivalent to the scaling of the full XCHF method. For the relatively small basis set used in the HCN calculation described in the next section, correlating two special electron orbitals to the proton orbital resulted in $\sim 3\times10^{10}$ integrals, or 230 GB of storage, while correlating four special electron orbitals would have required $\sim 9\times10^{12}$ integrals, or 70 TB of storage. We expect that the majority of these integrals will be very small, particularly when they involve geminal terms connecting electronic AOs that are far from the nuclear AOs, and can be successfully screened. These modifications to our integral code are currently in development and are crucial to the success of the method in the future. However, even with appropriate integral screening, direct SCF, which requires on-the-fly re-evaluation of the integrals at each SCF step, is likely not practical with currently available resources. Our current strategy requires that at least the four-particle integrals be stored in memory or on disk.

To this end, we have adopted a hybrid MPI/OpenMP protocol within our in-house NEO code. The integral calculation is embarrassingly parallelizable, as each integral can be calculated completely independently from the others, leading to almost ideal scaling with respect to the number of MPI processes or OpenMP threads. The use of this parallelization protocol not only enables the distribution of the calculation of the large number of integrals to many different cores, but it also obviates the need for the storage of all integrals on a single node because each MPI process need only store the integrals for which it is responsible. Steps such as the Fock matrix construction can be completed as reduction steps over intermediate Fock matrices calculated by each process.

Current directions for the integral code include adapting a Rys polynomial-based strategy for evaluation,[69] possibly in conjunction with RI approximations, as well as the introduction of

integral screening techniques. These developments, in conjunction with the existing parallel framework, will enable tractable calculations of larger systems of chemical interest.

8.4. Results and Discussion

In this section, we describe the application of the RXCHF methods to the hydrogen cyanide (HCN) molecule. We performed all NEO calculations using our in-house code and all non-NEO electronic structure calculations using the GAMESS package.[71] The heavy nuclei are fixed at their positions in the equilibrium geometry calculated at the RHF level using the pc0 electronic basis set,[72,73] and the centers of the proton basis functions are assigned to the position of the hydrogen at this equilibrium geometry. The single proton is treated quantum mechanically, and the fourteen electrons are partitioned into two special and twelve regular electrons (i.e., $N_s = 2$ and $N_r = 12$). The pc0 electronic basis set[72,73] was used for all atoms with an additional diffuse electronic basis function[74] placed on the H center. The nuclear basis set consisted of eleven AOs: an even-tempered set of five *s* functions[29] placed at the H center and two additional sets of three *s* functions, obtained from the same even-tempered set with the most diffuse and contracted functions removed, placed at positions ±0.2Å relative to the H center along the H-C-N axis. The extra basis function centers enable anisotropic nuclear density distributions. This basis set was developed for smaller test systems and adopted for the HCN calculation without any re-optimization of exponents or centers. The geminal parameters ($N_{\text{gem}} = 3$) were obtained from previous work,[75] in which they were variationally optimized at the XCHF/cc-pVDZ[76,77] level for the two-electron HeH^+ system and were fixed for all RXCHF calculations presented herein without re-optimization. These parameters are reported in Table 8.1. In our calculations, an

additional geminal factor with $b=1/2$ and $\gamma \approx 0$ is included to recover the $1+G$ ansatz as described in Section 8.2.

To benchmark the RXCHF method, we performed three-dimensional Fourier grid Hamiltonian[78] (FGH) calculations, in which a potential energy surface was generated at the RHF/pc0 level by moving the proton on a three-dimensional grid. The three-dimensional Schrödinger equation for the proton moving on this potential energy surface was then solved using the FGH method with 32 grid points in each dimension. The FGH approach is an electronically adiabatic method and therefore neglects nonadiabatic effects between the electrons and the quantum proton. However, nonadiabatic effects are not expected to be important for the HCN molecule, and thus the FGH method provides accurate data for benchmarking the RXCHF results.

An important property that the RXCHF methods must be able to accurately capture is the nuclear density, defined as

$$\rho^p(\mathbf{r}) = \frac{\left\langle \Psi^{\text{RXCHF-ne}} \left| \delta\left(\mathbf{r}-\mathbf{r}^p\right) \right| \Psi^{\text{RXCHF-ne}} \right\rangle}{\left\langle \Psi^{\text{RXCHF-ne}} \middle| \Psi^{\text{RXCHF-ne}} \right\rangle}. \tag{8.32}$$

As discussed previously, the RXCHF-ne form of the wavefunction is used even for the RXCHF-ae method since the latter method does not have a corresponding wavefunction. Providing accurate nuclear densities is critical to the applicability of this method for calculating rate constants, isotope effects and other properties for PCET systems. As discussed in the Introduction, mean-field-based approaches provide nuclear densities that are much too localized, resulting in non-physical molecular properties.

The nuclear densities obtained with the NEO-HF, NEO-RXCHF, and FGH methods are depicted in Figure 8.1. The NEO-HF densities are calculated using Eq. (8.32) with $\Psi^{\text{RXCHF-ne}}$

replaced by $\Phi^e\chi^p$. The origin was assigned to the position of the carbon with the H located on the positive *z*-axis. The RXCHF results are in excellent agreement with the FGH result, especially considering that we did not re-optimize the basis sets or geminal parameters. As expected, the NEO-HF method leads to overlocalized nuclear densities, a symptom of the lack of dynamical electron-proton correlation. To better quantify the comparison between these methods, Table 8.2 reports the characteristics of the nuclear density obtained by fitting the nuclear density to a one-dimensional Gaussian function, $\rho^{\text{fit}}(r) = A\exp\left(-\alpha\left(r - r_0\right)\right)$, where A is a normalization constant, r_0 is the position corresponding to the center of the Gaussian, and α is the exponent related to the frequencies in Table 8.2 through the relationship $\nu = \alpha/m_p$. The frequencies corresponding to the C-H stretch differ from the grid frequencies by ~70cm^{-1} which is considered to be in very good agreement for this method, especially in contrast the substantial error of ~2500 cm^{-1} associated with the NEO-HF frequency. The location of the maxima of the density profiles for the RXCHF methods differ by ~0.01 Å compared to the FGH maximum, which is also considered to be in reasonable agreement. We expect that the agreement would be improved through optimization of the electronic and nuclear basis sets, as well as the geminal parameters.

The objective of this paper is to illustrate that the NEO-RXCHF method provides qualitatively accurate nuclear densities. To date, other multicomponent molecular orbital methods have not achieved even qualitative accuracy.[22,23,30-32] Subsequent work will focus on improving the quantitative accuracy by optimizing the basis sets and geminal parameters. We emphasize that many standard quantum dynamical methods can provide more accurate nuclear densities for electronically adiabatic systems. The NEO-RXCHF method is designed to include non-Born-Oppenheimer effects that are not described by these standard methods. The

applications discussed herein are for benchmarking purposes only. Subsequently, the RXCHF methods will be applied to non-Born-Oppenheimer systems.

As discussed in Section 8.3, we were able to gain significant improvements in convergence of the RXCHF calculations on HCN when appropriate guess orbitals were used. In Figure 8.2, the converged occupied orbitals obtained from an RHF calculation using the same basis set as the RXCHF calculations are depicted in the top row. By inspection we can identify one orbital in the second column as a good candidate for a guess special MO because it is localized over the C-H bond. However, a more general procedure that is likely to be applicable to other systems would be to generate localized orbitals from the RHF orbitals. We thus applied the Boys localization procedure[65] to generate the localized orbitals depicted in the second row of Figure 8.2. Qualitatively, the orbital in the second column is slightly more localized than the RHF counterpart, and thus we proceed with this orbital as our guess for the single special electron MO. We used the remaining localized orbitals as a guess set for the regular electron MOs. This choice of guess orbitals worked well for these calculations, and we expect this strategy to provide optimal convergence for larger chemical systems.

Figure 8.2 also includes images of the converged regular and special electron MOs for the HCN molecule at the RXCHF-ne level. The resulting orbitals from the RXCHF-ae calculation appear qualitatively identical to this set and are included in Section G.6. As expected, the special electron MO from RXCHF is localized over the C-H bond, even more so than the most similar orbital from the RHF or localized orbital sets. Furthermore, the regular electron MOs are qualitatively similar to the RHF orbitals from the top row, as are the corresponding converged NEO-HF orbitals, which are provided in Section G.6. The similarity of the electronic MOs obtained with the RXCHF method to those obtained with the NEO-HF method, which

generates the relatively poor nuclear density profile in Figure 8.1, demonstrates the validity of the assumption underlying the RXCHF-ne ansatz. Specifically, coupling regular HF-like orbitals to a subsystem comprised of the explicitly-correlated special electron and nuclear MOs through a mean-field approach is sufficient to generate accurate nuclear densities. The failure of the NEO-HF method to generate even qualitatively reasonable nuclear densities demonstrates the need for explicit electron-proton correlation. Furthermore, these results indicate that sufficient accuracy can be attained without explicit correlation between the regular electron MOs and the nuclear MO and without exchange interactions between the regular and special electrons.

An additional encouraging result from the RXCHF calculations is that the special electron MO is predominantly localized over the C-H bond. We thus expect that strategies restricting the atomic basis set used to expand the special electron MOs in Eq. (8.29) could provide a promising direction to drastically improve the computational tractability of the RXCHF methods. Specifically, if the size of the special electron AO basis set is $N_{\mathrm{ebf}}^{s} << N_{\mathrm{ebf}}$, then the number of integrals to be evaluated reduces from $3N_{\mathrm{pbf}}^{2}N_{\mathrm{ebf}}^{6}$ to $3N_{\mathrm{pbf}}^{2}N_{\mathrm{ebf}}^{2}\left(N_{\mathrm{ebf}}^{s}\right)^{4}$ for $N_s \leq 2$ and from $2N_{\mathrm{pbf}}^{2}N_{\mathrm{ebf}}^{8}$ to $2N_{\mathrm{pbf}}^{2}N_{\mathrm{ebf}}^{2}\left(N_{\mathrm{ebf}}^{s}\right)^{6}$ for $N_s \geq 4$. As an example, for the HCN molecule, we envisage restricting the AO basis set of the single special electron MO to be comprised solely of valence *s* AOs on carbon and hydrogen, as well as *p* AOs aligned along the C-H axis. Eliminating the need for *p* AOs in the other Cartesian directions as well as any AOs on nitrogen will drastically improve the scaling of the RXCHF methods. Moreover, as shown by Figure 8.2 and the actual MO coefficients, these AOs do not contribute significantly to the special electron MO. We are currently investigating the adaptation of the RXCHF method to the use of a restricted basis set for the special electron MO. Certain complications arise when trying to maintain orthogonality between the regular and special electron MOs when a different AO

basis set is used for the latter, but these technical issues can be circumvented. A detailed presentation of this method as well as applications to a wide range of chemical systems will be presented in a future study.

8.5. Conclusions

In this paper, we extended the NEO-RXCHF approach to enable the study of systems in which a select proton is treated quantum mechanically. We derived the necessary equations to implement the NEO-RXCHF method and highlighted important technical considerations with respect to orthogonality, convergence, and parallelization. We also presented an application of the method to the hydrogen cyanide molecule and observed that the RXCHF results are in good agreement with grid-based calculations without any re-parameterization or re-optimization of basis sets or germinal parameters. The resulting RXCHF-ne and RXCHF-ae methods developed here are applicable to the general case of explicitly correlating any number of electronic MOs to the nuclear MO and will enable the study of a wide range of proton-containing systems. Although the formalism is presented for a single quantum proton, the extension to systems with multiple quantum protons is straightforward.[79] Moreover, the equations presented herein are also easily extended to systems with an odd number of electrons using analogs of unrestricted Hartree-Fock and restricted open-shell Hartree-Fock treatments.

We have also highlighted several promising directions for future research and further extensions of the RXCHF method described in this study. Most notably, developing a scheme to restrict the basis set of the special electron MO is physically warranted and would provide a significant increase in tractability, thereby enabling calculations on much larger molecules. Furthermore, enhancements of the integral code to include integral screening and more efficient

strategies for integral calculations will offer substantial gains in tractability. In addition, extensions of this methodology to open-shell systems, as well as to systems requiring multiconfigurational treatments, will allow the study of PCET systems such as the phenoxyl-phenol radical couple. Non-Born-Oppenheimer effects between the electrons and the transferring proton have been shown to be important in these types of systems.[5] The investigation of PCET with the NEO-RXCHF method will provide fundamental insights into the coupling between the electrons and the transferring proton and will enable the calculation of rate constants and kinetic isotope effects using existing PCET theories.

b_k	γ_k
7.6150	0.2148
5.0620	2.0946
2.9370	18.414

Table 8.1: Geminal parameters optimized at the XCHF level using the cc-pVDZ electronic basis set for the HeH^+ system taken from Ref. 75.

Method	ν (cm^{-1})	r_0 (Å)
NEO-HF	5077	1.052
RXCHF-ne	3604	1.046
RXCHF-ae	3476	1.041
3-D FGH	3544	1.056

Table 8.2: Characteristics of the nuclear density profiles depicted in Figure 8.1 when they are fit to a Gaussian with frequency ν and center r_0, as defined in the text.

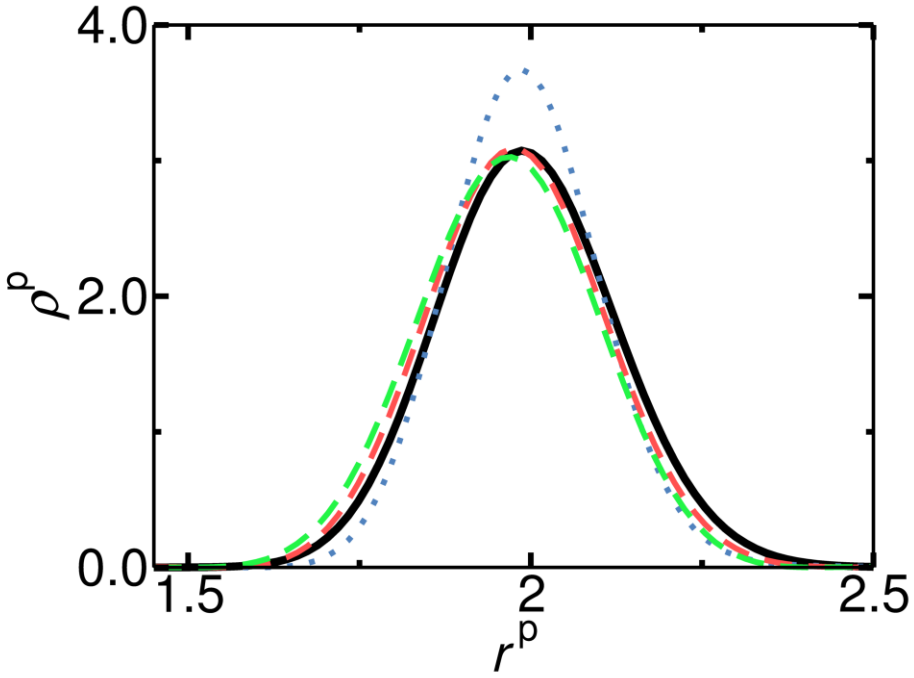

Figure 8.1: Nuclear density, ρ^p, corresponding to the quantum proton in HCN calculated along the HCN axis, r^p, using the 3D FGH (black, solid), NEO-HF (blue, dotted), RXCHF-ne (red, dashed) and RXCHF-ae (green, dashed) methods.

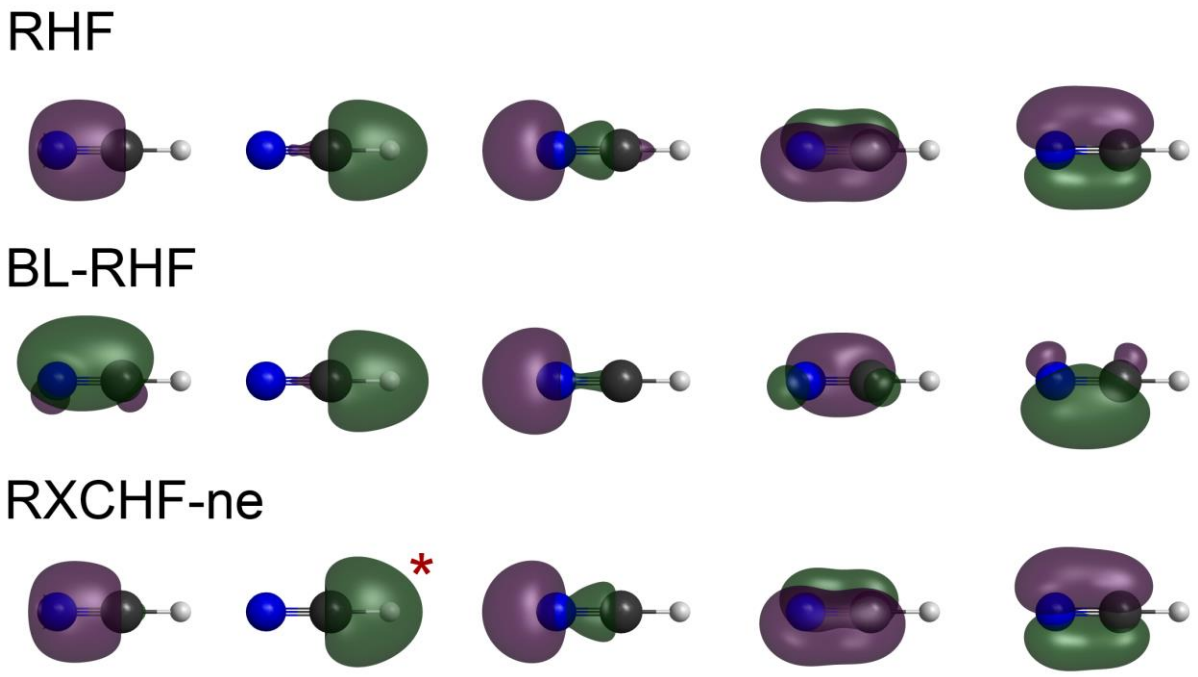

Figure 8.2: Electronic molecular orbitals for HCN obtained using different methods, where green and purple denote the different phases of the orbitals. The top row consists of five of the occupied MOs obtained from a converged RHF calculation ordered from left to right according to their respective eigenvalues from lowest to highest. The middle row consists of five of the Boys-localized orbitals generated from the RHF orbitals of the top row and are ordered to best match the character of the orbitals in the top row. The bottom row consists of four of the occupied regular MOs and the special MO obtained from the converged RXCHF-ne calculation. The regular MOs are ordered with respect to their respective eigenvalues from lowest to highest, but the special MO, denoted with an asterisk, is placed in between them to highlight its similarity to the MOs in the previous rows. In all rows, the two occupied or localized MOs that were excluded from this image correspond to the core orbitals on the N and C centers (i.e., the two lowest-eigenvalue orbitals for the RHF and RXCHF-ne calculations).

References

(1) Y. Cha, C. J. Murray, and J. P. Klinman, *Science* **243**, 1325 (1989).
(2) S. Raugei and M. L. Klein, *Journal of the American Chemical Society* **125**, 8992 (2003).
(3) M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello, *Science* **275**, 817 (1997).
(4) S. Hammes-Schiffer, *Journal of Physical Chemistry Letters* **2**, 1410 (2011).
(5) A. Sirjoosingh and S. Hammes-Schiffer, *Journal of Physical Chemistry A* **115**, 2367 (2011).
(6) J. H. Skone, A. V. Soudackov, and S. Hammes-Schiffer, *Journal of the American Chemical Society* **128**, 16655 (2006).
(7) R. I. Cukier and D. G. Nocera, *Annual Review of Physical Chemistry* **49**, 337 (1998).
(8) S. Hammes-Schiffer, *Chemical Reviews* **110**, 6937 (2010).
(9) D. R. Weinberg, C. J. Gagliardi, J. F. Hull, C. F. Murphy, C. A. Kent, B. C. Westlake, A. Paul, D. H. Ess, D. G. McCafferty, and T. J. Meyer, *Chemical Reviews* **112**, 4016 (2012).
(10) A. D. Bochevarov, E. F. Valeev, and C. D. Sherrill, *Molecular Physics* **102**, 111 (2004).
(11) S. Bubin, M. Pavanelo, W. C. Tung, K. L. Sharkey, and L. Adamowicz, *Chemical Reviews* **113**, 36 (2013).
(12) M. Cafiero, S. Bubin, and L. Adamowicz, *Physical Chemistry Chemical Physics* **5**, 1491 (2003).
(13) L. S. Cederbaum, *Journal of Chemical Physics* **138**, 224110 (2013).
(14) A. Chakraborty, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Physics* **129**, 014101 (2008).
(15) A. Chakraborty, M. V. Pak, and S. Hammes-Schiffer, *Physical Review Letters* **101**, 153001 (2008).
(16) M. Hoshino, H. Nishizawa, and H. Nakai, *Journal of Chemical Physics* **135**, 024111 (2011).
(17) E. Kamarchik and D. A. Mazziotti, *Physical Review A* **75**, 013203 (2007).
(18) E. Kamarchik and D. A. Mazziotti, *Physical Review A* **79**, 012502 (2009).
(19) T. Kreibich and E. K. U. Gross, *Physical Review Letters* **86**, 2984 (2001).
(20) T. Kreibich, R. van Leeuwen, and E. K. U. Gross, *Physical Review A* **78**, 022501 (2008).
(21) E. Matyus and M. Reiher, *Journal of Chemical Physics* **137**, 024104 (2012).
(22) H. Nakai, *International Journal of Quantum Chemistry* **86**, 511 (2002).
(23) H. Nakai, K. Sodeyama, and M. Hoshino, *Chemical Physics Letters* **345**, 118 (2001).
(24) Y. Shigeta, H. Nagao, K. Nishikawa, and K. Yamaguchi, *Journal of Chemical Physics* **111**, 6171 (1999).
(25) T. Shimazaki and M. Kubo, *Chemical Physics Letters* **525-26**, 134 (2012).
(26) A. Sirjoosingh, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Theory and Computation* **7**, 2689 (2011).
(27) A. Sirjoosingh, M. V. Pak, C. Swalina, and S. Hammes-Schiffer, *Journal of Chemical Physics* **139**, 034102 (2013).
(28) J. H. Skone, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Physics* **123**, 134108 (2005).
(29) C. Swalina, M. V. Pak, A. Chakraborty, and S. Hammes-Schiffer, *Journal of Physical Chemistry A* **110**, 9983 (2006).
(30) C. Swalina, M. V. Pak, and S. Hammes-Schiffer, *Chemical Physics Letters* **404**, 394 (2005).

(31) M. Tachikawa, K. Mori, H. Nakai, and K. Iguchi, *Chemical Physics Letters* **290**, 437 (1998).
(32) S. P. Webb, T. Iordanov, and S. Hammes-Schiffer, *Journal of Chemical Physics* **117**, 4106 (2002).
(33) S. Hammes-Schiffer and A. V. Soudackov, *Journal of Physical Chemistry B* **112**, 14108 (2008).
(34) M. V. Pak and S. Hammes-Schiffer, *Physical Review Letters* **92**, 103002 (2004).
(35) M. V. Pak, C. Swalina, S. P. Webb, and S. Hammes-Schiffer, *Chemical Physics* **304**, 227 (2004).
(36) C. Swalina and S. Hammes-Schiffer, *Journal of Physical Chemistry A* **109**, 10410 (2005).
(37) C. Ko, M. V. Pak, C. Swalina, and S. Hammes-Schiffer, *Journal of Chemical Physics* **135**, 054106 (2011).
(38) W. Cencek, J. Komasa, and J. Rychlewski, *Chemical Physics Letters* **246**, 417 (1995).
(39) W. Cencek and J. Rychlewski, *Journal of Chemical Physics* **98**, 1252 (1993).
(40) S. Kedzuch, M. Milko, and J. Noga, *International Journal of Quantum Chemistry* **105**, 929 (2005).
(41) W. Klopper and W. Kutzelnigg, *Journal of Chemical Physics* **94**, 2020 (1991).
(42) T. Korona, H. L. Williams, R. Bukowski, B. Jeziorski, and K. Szalewicz, *Journal of Chemical Physics* **106**, 5109 (1997).
(43) W. Kutzelnigg and W. Klopper, *Journal of Chemical Physics* **94**, 1985 (1991).
(44) F. R. Manby, H. J. Werner, T. B. Adler, and A. J. May, *Journal of Chemical Physics* **124** (2006).
(45) D. A. Mazziotti, *Journal of Chemical Physics* **112**, 10125 (2000).
(46) B. J. Persson and P. R. Taylor, *Journal of Chemical Physics* **105**, 5915 (1996).
(47) V. A. Rassolov, *Journal of Chemical Physics* **117**, 5978 (2002).
(48) S. Ten-no, *Chemical Physics Letters* **330**, 169 (2000).
(49) V. Termath, W. Klopper, and W. Kutzelnigg, *Journal of Chemical Physics* **94**, 2002 (1991).
(50) E. F. Valeev, *Chemical Physics Letters* **395**, 190 (2004).
(51) E. F. Valeev, *Journal of Chemical Physics* **125** (2006).
(52) S. A. Varganov and T. J. Martinez, *Journal of Chemical Physics* **132**, 054103 (2010).
(53) A. Sirjoosingh, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Physics* **136**, 174114 (2012).
(54) A. Chakraborty and S. Hammes-Schiffer, *Journal of Chemical Physics* **129**, 204101 (2008).
(55) A. Sirjoosingh, M. V. Pak, C. Swalina, and S. Hammes-Schiffer, *Journal of Chemical Physics* **139**, 034103 (2013).
(56) A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. (Dover, New York, 1996).
(57) F. W. Bobrowicz and W. A. Goddard, *Methods of Electronic Structure Theory*. (Plenum, New York, 1977).
(58) W. J. Hunt, T. H. Dunning, and W. A. Goddard, *Chemical Physics Letters* **3**, 606 (1969).
(59) W. A. Goddard, T. H. Dunning, and W. J. Hunt, *Chemical Physics Letters* **4**, 231 (1969).
(60) W. J. Hunt, W. A. Goddard, and T. H. Dunning, *Chemical Physics Letters* **6**, 147 (1970).
(61) G. Chaban, M. W. Schmidt, and M. S. Gordon, *Theoretical Chemistry Accounts* **97**, 88 (1997).

(62) J. Douady, Y. Ellinger, R. Subra, and B. Levy, *Journal of Chemical Physics* **72**, 1452 (1980).
(63) T. H. Fischer and J. Almlof, *Journal of Physical Chemistry* **96**, 9768 (1992).
(64) R. Fletcher, *Practical Methods of Optimization*. (Wiley, New York, 1980).
(65) J. M. Foster and S. F. Boys, *Reviews of Modern Physics* **32**, 300 (1960).
(66) C. Edmiston and K. Ruedenberg, *Reviews of Modern Physics* **35**, 457 (1963).
(67) L. E. McMurchie and E. R. Davidson, *Journal of Computational Physics* **26**, 218 (1978).
(68) B. J. Persson and P. R. Taylor, *Theoretical Chemistry Accounts* **97**, 240 (1997).
(69) A. Komornicki and H. F. King, *Journal of Chemical Physics* **134**, 244115 (2011).
(70) M. Dupuis, J. Rys, and H. F. King, *Journal of Chemical Physics* **65**, 111 (1976).
(71) M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, *Journal of Computational Chemistry* **14**, 1347 (1993).
(72) F. Jensen, *Journal of Chemical Physics* **115**, 9113 (2001).
(73) F. Jensen, *Journal of Chemical Physics* **116**, 7372 (2002).
(74) F. Jensen, *Journal of Chemical Physics* **117**, 9234 (2002).
(75) C. Swalina, *unpublished work* (2011).
(76) T. H. Dunning, *Journal of Chemical Physics* **90**, 1007 (1989).
(77) D. E. Woon and T. H. Dunning, *Journal of Chemical Physics* **100**, 2975 (1994).
(78) C. C. Marston and G. G. Balint-Kurti, *Journal of Chemical Physics* **91**, 3571 (1989).
(79) B. Auer and S. Hammes-Schiffer, *Journal of Chemical Physics* **132**, 084110 (2010).

Chapter 9

Conclusions

9.1. Summary

In this dissertation, we discussed several aspects of electron-proton nonadiabaticity, their implications for the study of proton-coupled electron transfer (PCET) reactions, and the development of electronic structure methods that account for these types of effects.

The introductory comments in Chapter 1 aimed to set the framework for the topics developed in this dissertation. In particular, we discussed the PCET class of reactions and gave several examples of the promising applications in which they play an integral role. We proceeded to discuss the importance of electron-proton nonadiabatic effects in these types of systems, and described the need for characterization of these types of effects in addition to the development of non-Born-Oppenheimer methods to be able to study these systems using theoretical approaches. We posited that the development of such methods would enable the application of computational and theoretical tools to PCET systems in order to provide fundamental insight into the physics of these reactions. Furthermore, such tools could be used to verify and predict experimentally relevant quantities such as rate constants and kinetic isotope effects for PCET reactions.

We first aimed to develop methodology that would enable the characterization of electron-proton nonadiabatic effects in PCET reactions. As discussed in Chapter 2, we developed several qualitative and quantitative diagnostics accessible from standard electronic structure calculations in order to characterize these types of effects. In particular, we related a high degree of electron-proton nonadiabaticity with the commensurate large degree of electronic charge redistribution during the proton motion, outlining qualitative pictures (electrostatic potential maps) and quantitative metrics (nonadiabatic couplings) as diagnostics which could be used to characterize electron-proton nonadiabatic effects. Furthermore, we developed additional quantitative diagnostics related to vibronic couplings, which are integral quantities used to calculate rate constants and kinetic isotope effects in previously developed PCET rate theories. When applied to the prototypical PCET reaction involving the phenoxyl/phenol radical couple, these diagnostics highlighted the importance of electron-proton nonadiabaticity in these types of systems, and this result set the stage for the development of non-Born-Oppenheimer methods presented in the later chapters of this dissertation.

An additional consideration that naturally arose in the study described in Chapter 2 was the development of diabatization schemes for PCET reactions. In particular, the previously developed PCET rate theories require, as a starting point, the definition of charge-localized reactant and product diabatic vibronic states which correspond to the transferring electron and proton localized on their respective donors and acceptors. In Chapter 2, and further extended in Chapter 3, we developed a straightforward scheme to generate such quantities for general PCET reactions and discussed the implications of the different variants of this procedure related to the application of the PCET rate theories previously discussed. The initial procedure generated charge-localized diabatic electronic states with respect to the motion of the transferring proton

and involved the application of an adiabatic-to-diabatic transformation, which generates rigorous diabatic states along one-dimensional coordinates associated with the motion of the transferring proton. We then described a subsequent step to generate charge-localized diabatic vibronic states which can be directly used as input quantities for existing PCET rate theories.

Returning to the development of non-Born-Oppenheimer electronic structure methods, the remainder of the dissertation described our progress in developing new methodology within the nuclear-electronic orbital (NEO) framework, where electrons and a subset of nuclei (i.e., the transferring proton in PCET systems) are treated quantum mechanically at the same level. This type of approach inherently includes nonadiabatic effects between electrons and the quantum nuclei and thus would be suitable for the study of PCET systems.

We first described a multicomponent density functional theory (MCDFT) approach within the nuclear-electronic orbital (NEO) framework in Chapters 4 and 5 where, analogous to electronic DFT, the ground state energy can be expressed as a functional of an electronic density and a density of the quantum nuclei. We described the derivation of new electron-proton correlation functionals defined in terms of pair densities corresponding to explicitly-correlated nuclear-electronic wavefunctions and demonstrated their effectiveness with suitable parameterization. Furthermore, in Chapter 5, we discussed the interplay between treating electron-electron and electron-proton correlation in model systems, and determined that these interactions were predominantly additive effects to the energy. This important conclusion implies that future development of electron-proton correlation functionals in the MCDFT approach can proceed essentially independently from the choice of electron-electron exchange-correlation functionals prevalent in electronic DFT, obviating the need for reparametrization of either when both are used in conjunction with each other.

Although Chapters 4 and 5 lay the framework for MCDFT methods that include electron-proton nonadiabatic effects, their current implementation requires the evaluation of pair densities of explicitly-correlated wavefunctions, which is computationally intensive for larger systems (e.g., for PCET systems). Further optimization of these functionals is a promising research direction and is discussed in Section 9.2.

Chapters 6-8 discuss the formulation and application of the reduced explicitly correlated Hartree-Fock (RXCHF) approach. This approach differs from previously developed explicitly-correlated NEO approaches in that a subset of electronic orbitals is correlated to the quantum nuclear orbital. The specific variants of NEO-RXCHF derived in Chapter 6 were for the case of explicitly correlating one electronic orbital, and we demonstrated that significant improvements in computational tractability could be obtained by making approximations to the electronic exchange contributions to the energy expression.

Applying the RXCHF methods to molecular positron-containing systems in Chapter 7 demonstrated the accuracy of the approximate-exchange variants in addition to their significantly higher degree of computational tractability as compared to the more expensive RXCHF approach with all exchange terms included. The positronic systems studied were highly nonadiabatic due to the light mass of the positron which was treated quantum mechanically on the same level as the electrons, and thus served as a good benchmark for the RXCHF method. In particular, RXCHF was able to accurately capture electronic and positronic densities and provided quantitatively accurate annihilation rates as compared to highly expensive benchmark methods. The continued application of RXCHF to these types of positron-containing systems is thus a promising direction of research as much larger molecular systems can be studied with this approach than could be studied with current methodology.

In Chapter 8, we extended the RXCHF method to the general case of any number of electronic orbitals explicitly correlated to the quantum nuclear orbital. This extension was necessary for the application of RXCHF to proton-containing systems as at least two electronic orbitals should be explicitly correlated to the quantum proton orbital. We demonstrated the quantitative accuracy of RXCHF when applied to an electronically adiabatic system, hydrogen cyanide, for which accurate benchmark data could be generated using standard Born-Oppenheimer approaches. In particular, the nuclear densities corresponding to the quantum proton generated by RXCHF were very accurate, even though approximations to electronic exchange terms were invoked in order to maintain a reasonable level of computational tractability. The success of the RXCHF method when applied to proton-containing systems is an exciting result as this work lays the foundation for future studies on PCET systems where electron-proton nonadiabatic effects are important and would be inherently included by RXCHF.

This dissertation clarified the importance of electron-proton nonadiabatic effects in PCET systems and developed diagnostics to characterize these effects. Furthermore, non-Born-Oppenheimer electronic structure methods were developed, and in addition to the MCDFT methods described, a new approach, NEO-RXCHF, was derived and demonstrated significant improvement in computational tractability while still maintaining accuracy when applied to positron- and proton-containing systems. Since this method includes electron-proton nonadiabatic effects, it can be applied to the study of PCET systems where fundamental insight can be gained into PCET mechanisms.

9.2. Future Directions

A very promising direction for the NEO-RXCHF method is the study of PCET systems. As Chapter 7 described the success of RXCHF when applied to the highly nonadiabatic

molecular positron-containing systems, and Chapter 8 described its potential for studying proton-containing systems, this dissertation provides the framework for the potentially powerful application of RXCHF to PCET systems. Being able to perform calculations on PCET systems such as the phenoxyl/phenol radical couple described in Chapter 2 using an *ab initio* electronic structure method that includes nonadiabatic effects is an exciting prospect, and this work is likely attainable in the near future. In particular, some of the strategies to further improve the computational tractability of RXCHF were discussed in Chapter 8 and are being currently investigated.

Furthermore, the use of RXCHF reference wavefunctions in multireference approaches, such as in a two-configuration self-consistent-field (TC-SCF) approach, will enable the direct study of PCET systems that require multiple configurations to describe donor-hydrogen and acceptor-hydrogen states. These methods are also accessible in the near future and are currently under development in our group. This methodology will enable the direct calculation of vibronic couplings using a first-principles approach which can be used in conjunction with PCET rate theories to determine experimentally relevant quantities such as rate constants and kinetic isotope effects.

Another interesting direction is the continued development of electron-proton correlation functionals for use in an MCDFT approach within the NEO framework. The use of accurate *ab initio* data generated from, for example, NEO-RXCHF, will enable reparameterization of the existing functionals for a representative test set of molecules. In addition, this benchmark data may provide insight into the development of new, more versatile functionals which could be developed while emphasizing tractability, as is often done in the development of electronic DFT functionals.

9.3. Concluding Remarks

This dissertation presented an analysis of electron-proton nonadiabaticity in the context of PCET reactions. We discussed methods to characterize these types of effects, and once determining their importance for general PCET processes, outlined the development of novel non-Born-Oppenheimer electronic structure methods within the NEO framework which treat electrons and select nuclei quantum mechanically. We described the development of electron-proton correlation functionals for use in MCDFT methods and outlined their prospects for future studies. Finally, we derived and provided a detailed analysis of the NEO-RXCHF method which explicitly correlates a subset of electronic orbitals to the quantum nuclear orbital. This method was very successful when applied to positron- and proton-containing species, and in conjunction with some of the future directions discussed in Section 9.2, provides an exciting framework within which accurate, yet tractable calculations can be performed on PCET systems. The application of this type of non-Born-Oppenheimer *ab initio* approach will enable an increased understanding of PCET mechanisms, offering insight and enabling calculations of experimentally relevant quantities for this important class of reactions.

Appendix A

Supporting Information for Chapter 2[‡]

A.1. Summary of Contents

In this Appendix, we include supporting information corresponding to the material presented in Chapter 2. We include a detailed analysis of the nonadiabatic coupling terms developed in Section 2.2. We also include the following figures: adiabatic and diabatic potential energy curves shown in Figure 2.1 plotted with diabatic electronic coupling; analog to Figure 2.7 for the second diabatic electronic state for phenoxyl-phenol system; analogs to Figures 2.7 and 2.8 depicting charge transfer properties of diabatic electronic states for benzyl-toluene system. We also include the following tables: parameters obtained from the diabatic electronic states to calculate the semiclassical vibronic coupling; a comparison of the nonadiabatic coupling terms in Eqs. (2.8) and (2.11) for the phenoxyl-phenol system.

[‡]Reproduced with permission from:
A. Sirjoosingh, and S. Hammes-Schiffer, "Proton-Coupled Electron Transfer versus Hydrogen Atom Transfer: Generation of Charge-Localized Diabatic States," *J. Phys. Chem. A* **115**, 2367-2377 (2011). © 2011 American Chemical Society

A.2. Analysis of Nonadiabatic Coupling Terms

Eq. (2.11) can be derived from Eq. (2.8) using the following expressions:

$$\begin{aligned}\left\langle \phi_\mu^{(i)} \left| \mathbf{d}_{ij}^{(ep)} \cdot \nabla_{\mathbf{r}_p} \phi_\nu^{(j)} \right\rangle_p &= \left\langle \mathbf{d}_{ij}^{(ep)} \cdot \nabla_{\mathbf{r}_p} \left(\phi_\nu^{(j)} \phi_\mu^{(i)} \right) \right\rangle - \left\langle \phi_\nu^{(j)} \left| \mathbf{d}_{ji}^{(ep)} \cdot \nabla_{\mathbf{r}_p} \phi_\mu^{(i)} \right\rangle_p \right. \\ &= -\left\langle \phi_\mu^{(i)} \left| \nabla_{\mathbf{r}_p} \cdot \mathbf{d}_{ij}^{(ep)} \right| \phi_\nu^{(j)} \right\rangle_p + \left\langle \phi_\nu^{(j)} \left| \mathbf{d}_{ji}^{(ep)} \cdot \nabla_{\mathbf{r}_p} \phi_\mu^{(i)} \right\rangle_p \right. \\ &= \left\langle \phi_\mu^{(i)} \left| f_{ij}^{(ep)} \right| \phi_\nu^{(j)} \right\rangle_p - \left\langle \phi_\mu^{(i)} \left| g_{ij}^{(ep)} \right| \phi_\nu^{(j)} \right\rangle_p + \left\langle \phi_\nu^{(j)} \left| \mathbf{d}_{ji}^{(ep)} \cdot \nabla_{\mathbf{r}_p} \phi_\mu^{(i)} \right\rangle_p \right.,\end{aligned} \tag{A.1}$$

where $f_{ij}^{(ep)} = \left\langle \nabla_{\mathbf{r}_p} \psi_i \cdot \nabla_{\mathbf{r}_p} \psi_j \right\rangle_e = \sum_k \mathbf{d}_{ki}^{(ep)} \cdot \mathbf{d}_{ki}^{(ep)}$ after insertion of the identity operator. (Note that a complete electronic basis set $\{\psi_k\}$ is assumed.) In Eq. (A.1), the first equality used the chain rule for $\nabla_{\mathbf{r}_p} \left(\phi_\nu^{(j)} \phi_\mu^{(i)} \right)$, the second equality used integration by parts (neglecting the term with the proton wavefunction evaluated at infinity) and the identity $\mathbf{d}_{ij}^{(ep)} = -\mathbf{d}_{ji}^{(ep)}$, and the third equality used the chain rule for $\nabla_{\mathbf{r}_p} \cdot \mathbf{d}_{ij}^{(ep)}$. Rearrangement of Eq. (A.1) leads to:

$$\left\langle \phi_\mu^{(i)} \left| g_{ij}^{(ep)} \right| \phi_\nu^{(j)} \right\rangle_p = -\left\langle \phi_\mu^{(i)} \left| \mathbf{d}_{ij}^{(ep)} \cdot \nabla_{\mathbf{r}_p} \phi_\nu^{(j)} \right\rangle_p + \left\langle \phi_\nu^{(j)} \left| \mathbf{d}_{ji}^{(ep)} \cdot \nabla_{\mathbf{r}_p} \phi_\mu^{(i)} \right\rangle_p - \left\langle \phi_\mu^{(i)} \left| f_{ij}^{(ep)} \right| \phi_\nu^{(j)} \right\rangle_p . \tag{A.2}$$

Substituting Eq. (A.2) into Eq. (2.8) leads to Eq. (2.11).

A common implementation of Eq. (2.8) is to assume that the second-order nonadiabatic coupling terms (i.e., those involving $g_{ij}^{(ep)}$) are negligible in comparison to the first-order nonadiabatic coupling terms (i.e., those involving $\mathbf{d}_{ij}^{(ep)}$). This approach may lead to a vibronic Hamiltonian that is not Hermitian (i.e., the vibronic Hamiltonian matrix is not symmetric) because, in general, $\left\langle \phi_\mu^{(i)} \left| \mathbf{d}_{ij}^{(ep)} \cdot \nabla_{\mathbf{r}_p} \phi_\nu^{(j)} \right\rangle_p \neq \left\langle \phi_\nu^{(j)} \left| \mathbf{d}_{ji}^{(ep)} \cdot \nabla_{\mathbf{r}_p} \phi_\mu^{(i)} \right\rangle_p \right.$. In contrast, the implementation of Eq. (2.11), in which the $\mathbf{d}_{ki}^{(ep)} \cdot \mathbf{d}_{kj}^{(ep)}$ terms are neglected for $k > \max\{i, j\}$, still retains the

symmetry of the vibronic Hamiltonian matrix. Further insight into the relative magnitudes of the terms in Eqs. (2.8) and (2.11) may be obtained from Eq. (A.2). For example, the identity $g_{ii}^{(ep)} = -\sum_{k} \left| \mathbf{d}_{ki}^{(ep)} \right|^2$ leads to the relation $\left| g_{ii}^{(ep)} \right| \geq \left| \mathbf{d}_{12}^{(ep)} \right|^2$, $i = 1, 2$. Note that the $\left| \mathbf{d}_{12}^{(ep)} \right|^2$ terms are included in the implementation of Eq. (2.11).

We have analyzed the nonadiabatic coupling terms in Eqs. (2.8) and (2.11) for the phenoxyl-phenol system along the one-dimensional proton coordinate. The results for matrix elements pertaining to the lowest two electronic states, $i, j \leq 2$, are presented in Tables A.2 and A.3 of Supporting Information. Table A.2 illustrates the asymmetry of the vibronic Hamiltonian matrix defined by Eq. (2.8) when the second-order nonadiabatic coupling terms are neglected. This table also indicates that the magnitudes of the second-order nonadiabatic coupling terms may be comparable to the magnitudes of the first-order nonadiabatic coupling terms in Eq. (2.8). Table A.3 illustrates that the $d_{ki}^{(ep)} d_{kj}^{(ep)}$ terms for $k > 2$ are much smaller than the other contributing terms in Eq. (2.11). Note that these higher-order terms can be included to achieve the desired level of convergence in the implementation of Eq. (2.11) without a substantial amount of additional effort. Overall, this example indicates that the implementation of Eq. (2.11), in which the $d_{ki}^{(ep)} d_{kj}^{(ep)}$ terms for $k > 2$ are neglected, is more accurate than the implementation of Eq. (2.8), in which the second-order nonadiabatic coupling terms $g_{ij}^{(ep)}$ are neglected.

A.3. Supporting Figures

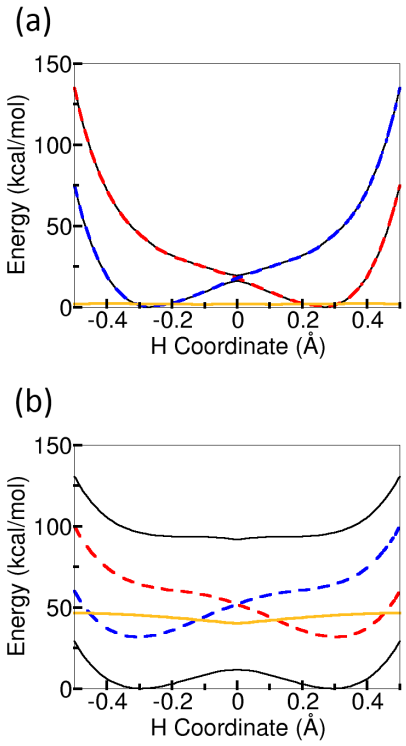

Figure A.1: Electronically adiabatic and diabatic potential energy curves as functions of the hydrogen coordinate for the (a) phenoxyl-phenol and (b) benzyl-toluene systems. The solid black curves are the ground and first excited state adiabatic energies $\varepsilon_1(r_p,\mathbf{R})$ and $\varepsilon_2(r_p,\mathbf{R})$, respectively, calculated with the CASSCF method. The dashed blue and red curves are the diabatic electronic energies $W_{11}(r_p,\mathbf{R})$ and $W_{22}(r_p,\mathbf{R})$, respectively, and the solid yellow curves are the diabatic electronic couplings $W_{12}(r_p,\mathbf{R})$ calculated from the expressions in Eq. (2.19) with $\gamma(r_0=0)=-\pi/4$.

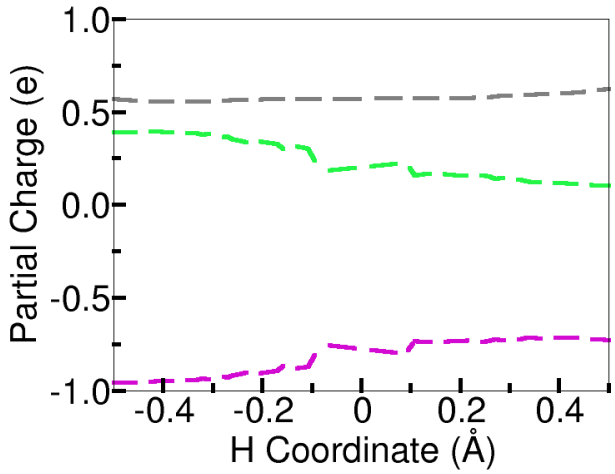

Figure A.2: Partial charges determined from electrostatic potential-derived atomic charges for the diabatic electronic state ξ_2 of the phenoxyl-phenol system. Partial charges are shown for the donor molecule (green), acceptor molecule (purple), and transferring hydrogen (grey). Calculated values of the partial charges around $r_p = 0$ are omitted due to numerical noise in this region.

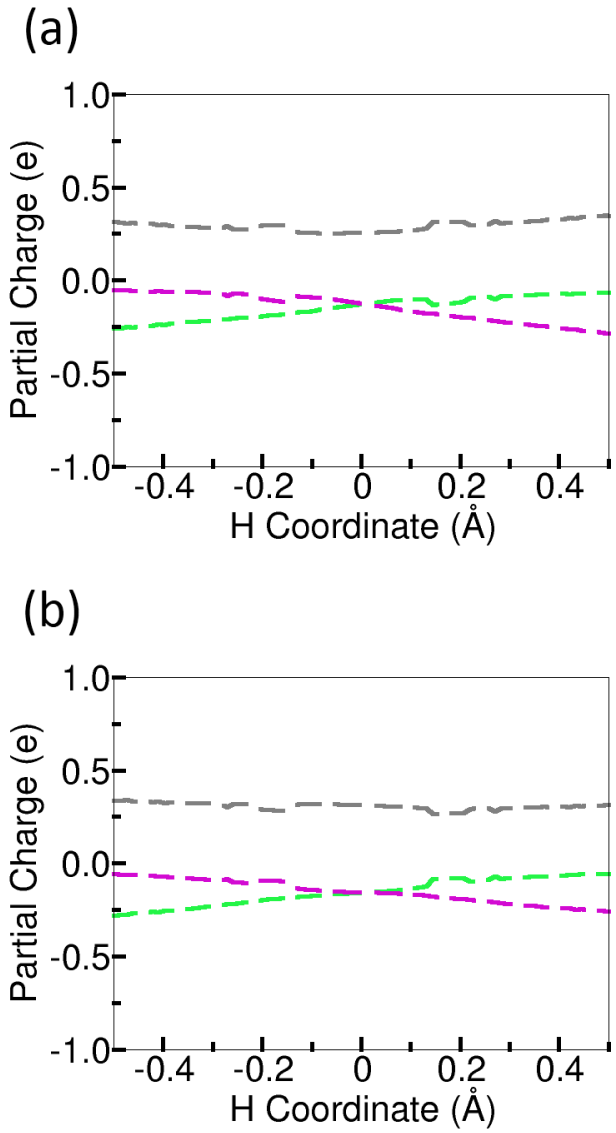

Figure A.3: Partial charges determined from electrostatic potential-derived atomic charges for the diabatic electronic states (a) ξ_1 and (b) ξ_2 of the benzyl-toluene system. Partial charges are shown for the donor molecule (green), acceptor molecule (purple), and transferring hydrogen (grey). Calculated values of the partial charges around $r_p = 0$ are omitted due to numerical noise in this region.

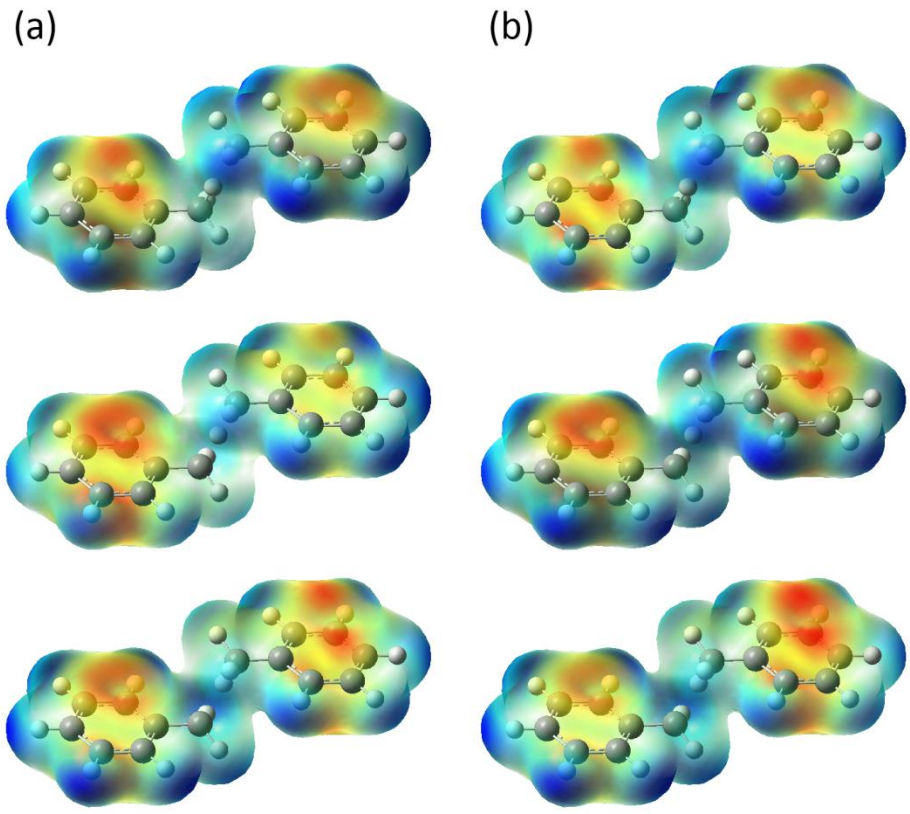

Figure A.4: Electrostatic potential maps for the diabatic electronic states (a) ξ_1 and (b) ξ_2 corresponding to a density isosurface value of 0.005 for the reactant (top), transition state (middle), and product (bottom) positions of the transferring hydrogen for the benzyl-toluene system. Negatively and positively charged regions are indicated by red and blue coloring, respectively.

A.4. Supporting Tables

System	p	κ	$V_{DA}^{(ad)}$ (cm^{-1})	$V_{DA}^{(sc)}$ (cm^{-1})
Phenol	0.013	0.265	27.6	7.3
Toluene	3.78	0.978	21.0	20.6

Table A.1: Parameters determined from the diabatic electronic states for the semiclassical calculation of vibronic couplings. For this calculation, the electronic coupling was approximated as a constant $W_{12}(r_p) = W_{12}(0)$ for all r_p.

(μ,ν)	$\langle d_{12}\rangle_{\mu\nu}$	$\langle d_{21}\rangle_{\nu\mu}$	$\langle g_{12}\rangle_{\mu\nu}$	$\langle g_{21}\rangle_{\nu\mu}$	$\langle g_{11}\rangle_{\mu\nu}$	$\langle g_{22}\rangle_{\mu\nu}$
(1,1)	-4.16×10^{-8}	-2.57×10^{-7}	-1.07×10^{-7}	1.09×10^{-7}	1.75×10^{-4}	6.84×10^{-3}
(1,2)	-1.67×10^{-3}	6.23×10^{-4}	1.15×10^{-3}	-1.15×10^{-3}	1.68×10^{-8}	1.37×10^{-8}
(1,3)	6.02×10^{-8}	1.08×10^{-7}	2.65×10^{-8}	-2.14×10^{-8}	-6.70×10^{-4}	-4.18×10^{-3}
(1,4)	-1.45×10^{-3}	4.30×10^{-4}	9.39×10^{-4}	-9.39×10^{-4}	2.35×10^{-9}	2.17×10^{-8}
(2,2)	-1.78×10^{-7}	6.89×10^{-8}	1.25×10^{-7}	-1.21×10^{-7}	1.74×10^{-5}	3.17×10^{-4}
(2,3)	-6.70×10^{-4}	-1.03×10^{-3}	-1.81×10^{-4}	1.81×10^{-4}	-7.08×10^{-8}	-7.14×10^{-9}
(2,4)	-1.50×10^{-7}	4.96×10^{-8}	9.75×10^{-8}	-1.03×10^{-7}	4.35×10^{-5}	3.47×10^{-4}
(3,3)	1.58×10^{-8}	-3.04×10^{-8}	-2.25×10^{-8}	2.37×10^{-8}	2.62×10^{-3}	2.61×10^{-3}
(3,4)	6.82×10^{-3}	-3.12×10^{-4}	-3.57×10^{-3}	3.57×10^{-3}	-2.58×10^{-8}	-1.21×10^{-8}

Table A.2: Nonadiabatic coupling terms[a,b] in the vibronic Hamiltonian matrix elements given by Eq. (2.8) for the phenoxyl-phenol system. The second-order nonadiabatic coupling terms are calculated using Eq. (A.2), where k ranges from 1 to 4 in the summation within the last term. Values are given in atomic units of energy.

[a] $\langle d_{ij}\rangle_{\mu\nu} \equiv -\frac{\hbar^2}{m_p}\left\langle \phi_\mu^{(i)} \middle| d_{ij}^{(ep)} \middle| \partial\phi_\nu^{(j)}/\partial r_p \right\rangle_p$

[b] $\langle g_{ij}\rangle_{\mu\nu} \equiv -\frac{\hbar^2}{2m_p}\left\langle \phi_\mu^{(i)} \middle| g_{ij}^{(ep)} \middle| \phi_\nu^{(j)} \right\rangle_p$

(μ,ν)	$\left\langle \bar{d}_{12} \right\rangle_{\mu\nu}$	$\left\langle d_{k1} d_{k2} \right\rangle_{\mu\nu}^{k=3,4}$	$\left\langle d_{k1} d_{k1} \right\rangle_{\mu\nu}^{k=1,2}$	$\left\langle d_{k1} d_{k1} \right\rangle_{\mu\nu}^{k=3,4}$	$\left\langle d_{k2} d_{k2} \right\rangle_{\mu\nu}^{k=1,2}$	$\left\langle d_{k2} d_{k2} \right\rangle_{\mu\nu}^{k=3,4}$
(1,1)	-1.49×10^{-7}	1.33×10^{-9}	-1.75×10^{-4}	8.25×10^{-11}	-6.84×10^{-3}	3.80×10^{-8}
(1,2)	-5.26×10^{-4}	-3.88×10^{-13}	-1.68×10^{-8}	-1.02×10^{-15}	-1.37×10^{-8}	1.94×10^{-12}
(1,3)	8.42×10^{-8}	2.55×10^{-9}	6.70×10^{-4}	-6.19×10^{-11}	4.18×10^{-3}	5.20×10^{-8}
(1,4)	-5.09×10^{-4}	5.83×10^{-13}	-2.35×10^{-9}	-3.47×10^{-15}	-2.17×10^{-8}	-1.68×10^{-12}
(2,2)	-5.43×10^{-8}	2.15×10^{-9}	-1.74×10^{-5}	6.96×10^{-11}	-3.17×10^{-4}	1.25×10^{-7}
(2,3)	-8.50×10^{-4}	-2.10×10^{-14}	7.08×10^{-8}	-4.48×10^{-15}	7.14×10^{-9}	6.19×10^{-12}
(2,4)	-5.05×10^{-8}	-2.53×10^{-9}	-4.35×10^{-5}	1.33×10^{-11}	-3.47×10^{-4}	-1.43×10^{-7}
(3,3)	-7.30×10^{-9}	5.65×10^{-10}	-2.62×10^{-3}	2.88×10^{-10}	-2.61×10^{-3}	2.56×10^{-7}
(3,4)	3.26×10^{-3}	6.03×10^{-14}	2.58×10^{-8}	1.10×10^{-15}	1.21×10^{-8}	-8.56×10^{-12}

Table A.3: Nonadiabatic coupling terms[a,b] in the vibronic Hamiltonian matrix elements given by Eq. (2.11) for the phenoxyl-phenol system. Note that $\left\langle \bar{d}_{11} \right\rangle_{\mu\nu}$, $\left\langle \bar{d}_{22} \right\rangle_{\mu\nu}$, and $\left\langle d_{k1} d_{k2} \right\rangle_{\mu\nu}^{k=1,2}$ are zero for all (μ,ν) pairs. The $d_{3i}^{(ep)}$ and $d_{4i}^{(ep)}$ terms were calculated from CASSCF(3,6) calculations with state averaging over four states with weights of 0.49, 0.49, 0.1 and 0.1 for the ground, first excited, second excited, and third excited electronic states, respectively. Values are given in atomic units of energy.

[a] $$\left\langle \bar{d}_{ij} \right\rangle_{\mu\nu} \equiv -\frac{\hbar^2}{2m_p}\left[\left\langle \phi_\mu^{(i)} \middle| d_{ij}^{(ep)} \middle| \partial\phi_\nu^{(j)} / \partial r_p \right\rangle_p + \left\langle \phi_\nu^{(j)} \middle| d_{ji}^{(ep)} \middle| \partial\phi_\mu^{(i)} / \partial r_p \right\rangle_p\right]$$

[b] $$\left\langle d_{ki} d_{kj} \right\rangle_{\mu\nu}^{k=a,b} \equiv \frac{\hbar^2}{2m_p}\left\langle \phi_\mu^{(i)} \middle| \sum_{k=a}^{b}\left(d_{ki}^{(ep)} d_{kj}^{(ep)}\right) \middle| \phi_\nu^{(j)} \right\rangle_p$$

Appendix B

Supporting Information for Chapter 3[‡]

B.1. Summary of Contents

In this Appendix, we include supporting information corresponding to the material presented in Chapter 3. We include the following figures: electrostatic potential maps corresponding to the diabatic electronic states at different positions of the transferring hydrogen for all three systems studied; partial charges of the donor molecule, acceptor molecule, and transferring hydrogen for the adiabatic and diabatic electronic states of the phenoxyl-quinol system and of the amidinium-carboxylate system prior to subtraction of the corresponding charges for the neutral system; adiabatic and diabatic electronic state energies along a normal mode coordinate and the IRC for the phenoxyl-quinol system. We also include the following table: average partial charges of the donor molecule, acceptor molecule, and transferring hydrogen for the three-dimensional adiabatic and diabatic electronic states of the phenoxyl-phenol system using the alternative more approximate approach. We also include the coordinates of the transition state geometries of the phenoxyl-quinol and amidinium-carboxylate systems.

[‡]Reproduced with permission from:

A. Sirjoosingh, and S. Hammes-Schiffer, "Diabatization Schemes for Generating Charge-Localized Electron-Proton Vibronic States in Proton-Coupled Electron Transfer Systems," *J. Chem. Theory Comput.* **7**, 2831-2841 (2011).

B.2. Supporting Figures

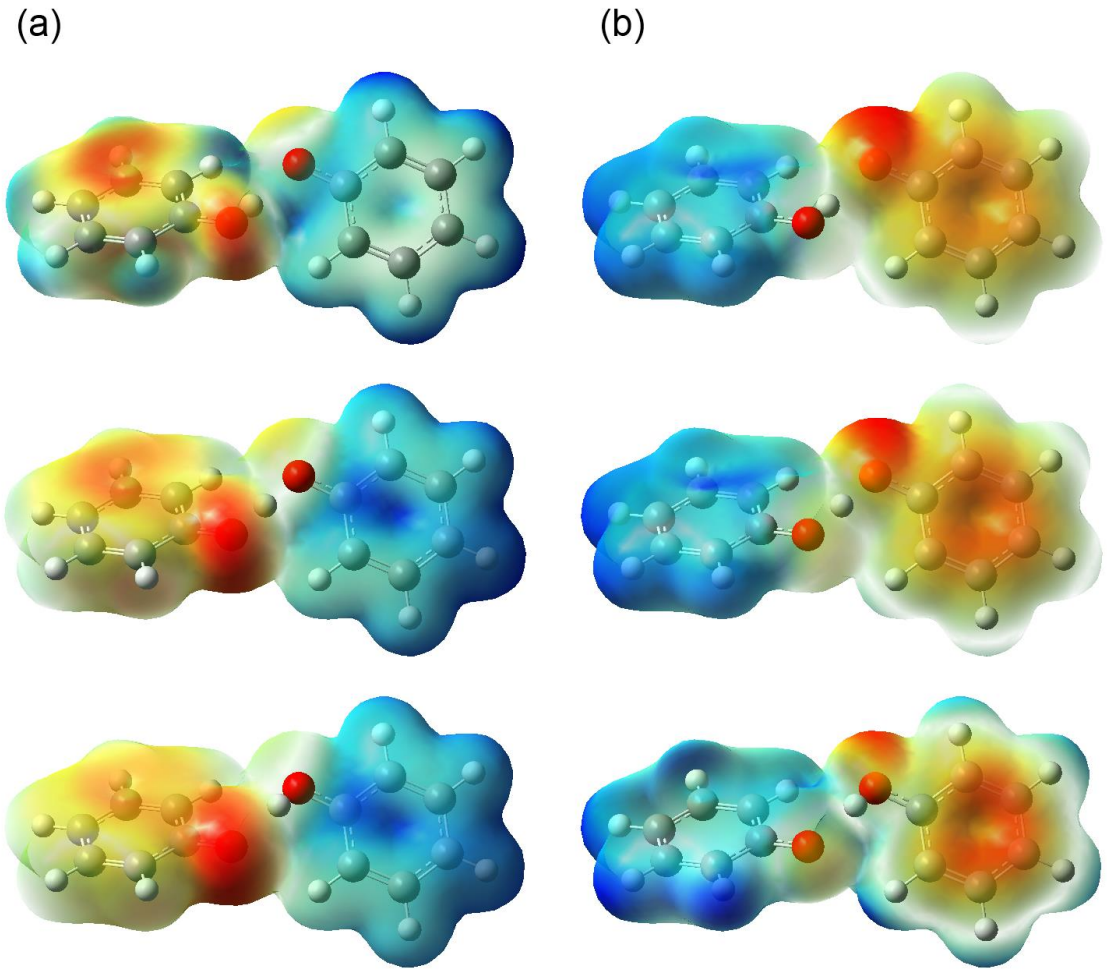

Figure B.1: Electrostatic potential maps for the diabatic electronic states (a) ξ_1 and (b) ξ_2 generated along the one-dimensional hydrogen coordinate corresponding to a density isosurface value of 0.005 for the reactant (top), transition state (middle), and product (bottom) positions of the transferring hydrogen for the phenoxyl-phenol system. Negatively and positively charged regions are indicated by red and blue coloring, respectively.

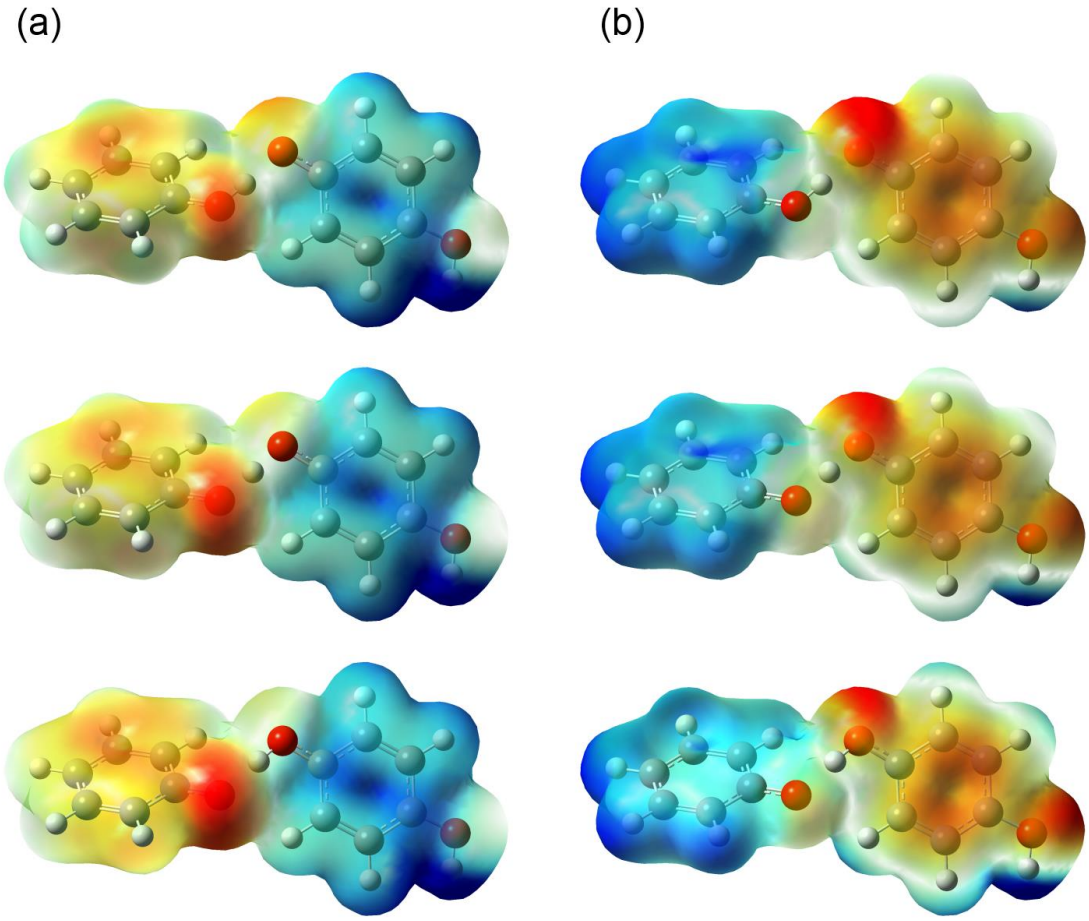

Figure B.2: Electrostatic potential maps for the diabatic electronic states (a) ξ_1 and (b) ξ_2 generated along the one-dimensional hydrogen coordinate corresponding to a density isosurface value of 0.005 for the reactant (top), transition state (middle), and product (bottom) positions of the transferring hydrogen for the phenoxyl-quinol system. Negatively and positively charged regions are indicated by red and blue coloring, respectively.

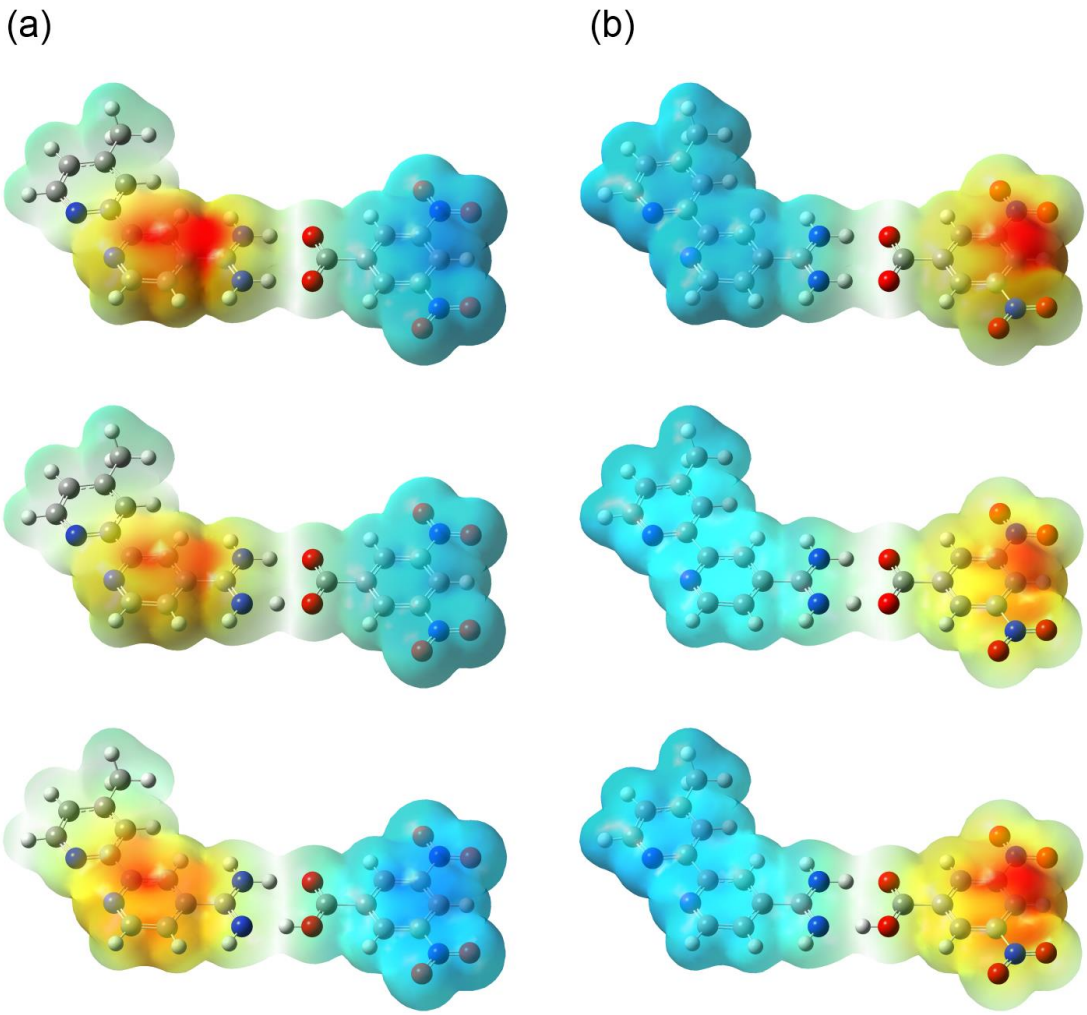

Figure B.3: Electrostatic potential maps for the diabatic electronic states (a) ξ_1 and (b) ξ_2 generated along the one-dimensional hydrogen coordinate corresponding to a density isosurface value of 0.005 for the reactant (top), midpoint (middle), and product (bottom) positions of the transferring hydrogen for the amidinium-carboxylate system. Negatively and positively charged regions are indicated by red and blue coloring, respectively. The maps for the amidinium-carboxylate system are difference electrostatic potential maps with respect to the neutral complex, as described in the text.

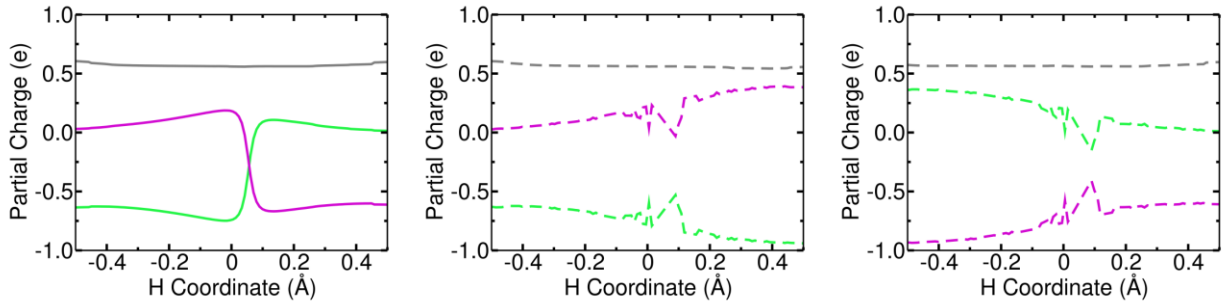

Figure B.4: Partial charges determined from electrostatic potential-derived atomic charges for the ground adiabatic electronic state (left), the diabatic electronic state ξ_1 (center) and the diabatic electronic state ξ_2 (right) as functions of the transferring hydrogen coordinate for the phenoxyl-quinol system. Partial charges are shown for the donor molecule (green), acceptor molecule (purple), and transferring hydrogen (gray). Calculated values of the partial charges for the diabatic electronic states around $r_p = r_0$ are omitted due to numerical noise in this region.

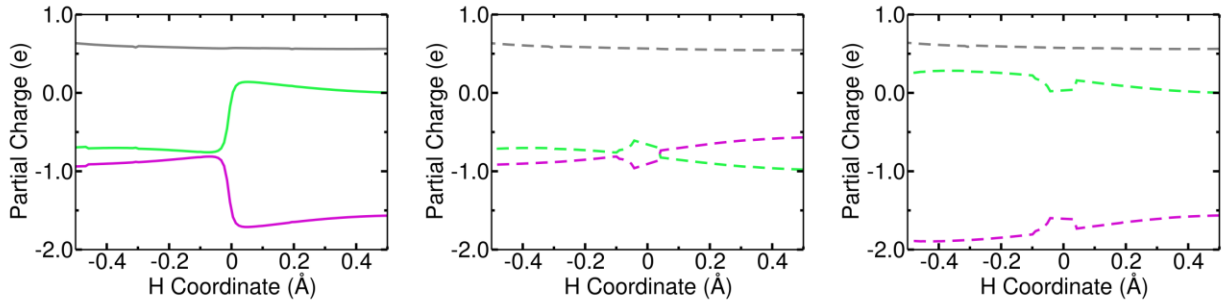

Figure B.5: Partial charges determined from electrostatic potential-derived atomic charges for the ground adiabatic electronic state (left), the diabatic electronic state ξ_1 (center) and the diabatic electronic state ξ_2 (right) as functions of the transferring hydrogen coordinate for the amidinium-carboxylate system. Partial charges are shown for the donor molecule (green), acceptor molecule (purple), and transferring hydrogen (gray). Calculated values of the partial charges for the diabatic electronic states around $r_p = r_0$ are omitted due to numerical noise in this region.

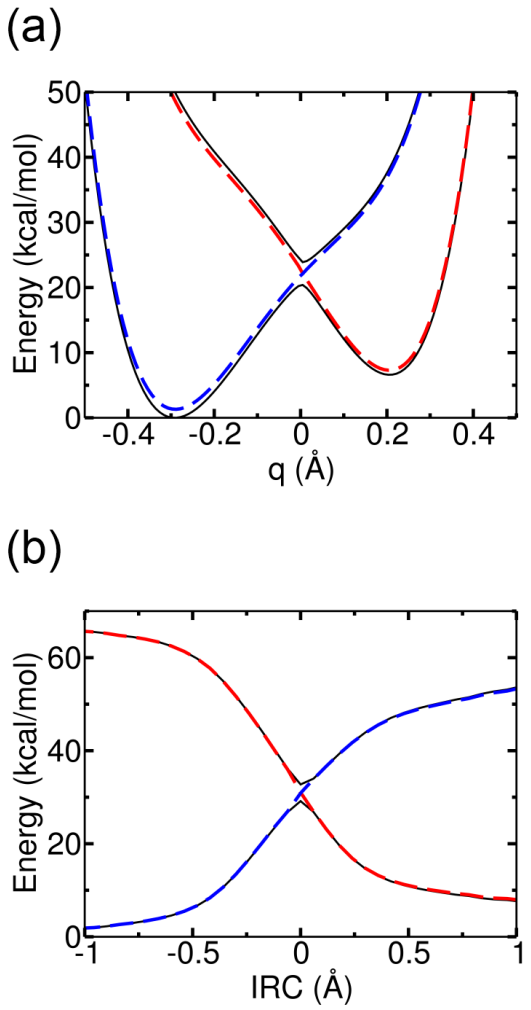

Figure B.6: Electronically adiabatic and diabatic potential energy curves for the phenoxyl-quinol system as functions of (a) the normal mode coordinate corresponding to the negative frequency at the transition state geometry and (b) the intrinsic reaction coordinate. The solid black curves are the ground and first excited adiabatic state energies $\varepsilon_1\left(r_p,\mathbf{R}\right)$ and $\varepsilon_2\left(r_p,\mathbf{R}\right)$, respectively, calculated with the CASSCF method, and the dashed blue and red curves are the diabatic electronic energies $W_{11}\left(r_p,\mathbf{R}\right)$ and $W_{22}\left(r_p,\mathbf{R}\right)$, respectively.

B.3. Supporting Tables

Electronic State	Donor Charge	Acceptor Charge	H Charge
ψ_1	–0.2±0.4	–0.2±0.4	0.5±0.1
ξ_1	–0.7±0.2	0.3±0.2	0.5±0.1
ξ_2	0.3±0.2	–0.7±0.2	0.5±0.1

Table B.1: Average electrostatic potential-derived partial charges calculated for the three-dimensional ground adiabatic and diabatic electronic states reported with standard deviations for the phenoxyl-phenol system. The tabulated results were computed by averaging over the 16^3 hydrogen positions on the three-dimensional grid, where points around $r_p = r_0$ were omitted due to numerical noise in this region. All charges are given in units of e. Deviation of the total charge from zero is due to numerical noise and round-off error. The diabatic electronic states were calculated using the more approximate formulation, where the adiabatic-to-diabatic transformation matrix depends explicitly on only the one-dimensional proton coordinate r_p. In this case, the first-order nonadiabatic couplings with respect to r_p vanish only along the one-dimensional proton donor-acceptor axis.

B.4. Transition State Geometries

Atomic coordinates for the optimized geometry of the phenoxyl-quinol system (Å)

C	3.56820200	-1.16380300	-0.63369600
C	2.44558900	-0.32288400	-0.36812300
C	2.63824100	0.80849800	0.48392500
C	3.88857700	1.07881700	1.01647600
C	4.98079400	0.24159300	0.74251900
C	4.80657100	-0.88050700	-0.08419500
O	1.29029400	-0.60993600	-0.89971400
O	-0.40350200	1.09608000	-0.73758100
C	-1.56887800	0.60980900	-0.39061400
C	-2.72680300	1.43383500	-0.50001600
C	-3.97331900	0.95790700	-0.15035400
C	-4.11509000	-0.36227100	0.32052200
C	-2.98658500	-1.19245400	0.44071200
C	-1.73168000	-0.71855600	0.10174800
H	-2.59569000	2.44368700	-0.87505200
H	-4.86009500	1.57833800	-0.23319600
O	-5.36870200	-0.77311800	0.65166500
H	-3.10153200	-2.20827400	0.81532600
H	-0.85359600	-1.34827600	0.19976100
H	1.78571500	1.44509700	0.70053000
H	4.02158000	1.94436600	1.66063900
H	5.95491700	0.45729100	1.17174300
H	5.65241700	-1.52936200	-0.29705400
H	3.41345200	-2.02012400	-1.28287800
H	0.42115400	0.28963900	-0.79755800
H	-5.33674600	-1.69648800	0.94935600

Atomic coordinates for the optimized geometry of the amidinium-carboxylate system (Å)

C	0.00000000	1.82352900	-0.67725500
N	0.00000000	1.02870600	0.35668900
N	0.00000000	1.30809600	-1.87475500
H	0.00000000	1.39120300	1.27725600
H	0.00000000	0.00000000	0.23942400
H	0.00000000	1.88690800	-2.67769600
H	0.00000000	0.28272700	-2.01207000
C	0.00000000	3.31448900	-0.49643800
C	0.00000000	3.88298600	0.76860200
C	0.00000000	5.26917200	0.91017300
N	0.00000000	6.06254000	-0.14715000
C	0.00000000	5.53215100	-1.34825000
C	0.00000000	4.16941400	-1.58752500
H	0.00000000	3.27158000	1.64446900
C	0.00000000	5.93577300	2.25809800
H	0.00000000	6.22127300	-2.17421800
H	0.00000000	3.82914400	-2.60404400
N	0.00000000	7.25578200	2.24853800
C	0.00000000	7.89232300	3.40194000
C	0.00000000	7.25887900	4.62994100
C	0.00000000	5.87216900	4.66275100
C	0.00000000	5.20868900	3.44521800
H	0.00000000	8.96740300	3.34762700
H	0.00000000	7.83655400	5.53692800
C	0.00000000	5.12311400	5.97177500
H	0.00000000	4.13637000	3.44702700
H	0.00000000	4.05102700	5.81727700
H	0.87569800	5.37913600	6.55922800
H	-0.87569800	5.37913600	6.55922800
C	0.00000000	-3.55838600	-1.33195200
C	0.00000000	-4.10413800	-2.60591800
C	0.00000000	-5.48029800	-2.75183900
C	0.00000000	-6.33850700	-1.67068200
C	0.00000000	-5.76508500	-0.41516700
C	0.00000000	-4.39408000	-0.22646900
H	0.00000000	-3.45896700	-3.46007400
N	0.00000000	-6.05308000	-4.09236000
H	0.00000000	-7.40003500	-1.80011000
N	0.00000000	-6.64274200	0.74874900
H	0.00000000	-3.97291800	0.75758000
O	0.00000000	-7.24319800	-4.18495400
O	0.00000000	-5.29804500	-5.01573200

O	0.00000000	-6.13129800	1.82642900
O	0.00000000	-7.82029600	0.55327200
C	0.00000000	-2.05402900	-1.14843300
O	0.00000000	-1.37201300	-2.18033500
O	0.00000000	-1.64073300	0.01791500

Appendix C

Supporting Information for Chapter 4[‡]

C.1. Summary of Contents

In this Appendix, we include supporting information corresponding to the material presented in Chapter 4. We include the following table: analog to Table 4.1 for the previously derived electron-proton functional.[1]

[‡]Reproduced with permission from:
A. Sirjoosingh, M. V. Pak, and S. Hammes-Schiffer, "Derivation of an Electron-Proton Correlation Functional for Multicomponent Density Functional Theory within the Nuclear-Electronic Orbital Approach," *J. Chem. Theory Comput.* **7**, 2689-2693 (2011). © 2011 American Chemical Society

C.2. Supporting Tables

Isotope	cc-pVDZ			cc-pVTZ		
	NEO-HF	NEO-DFT	FGH	NEO-HF	NEO-DFT	FGH
H	3098	1191	1191	3122	1098	1111
D	2284	820	805	2330	761	740
T	1903	660	640	1954	618	581

Table C.1: Vibrational frequencies in cm^{-1} corresponding to the hydrogen vibrational stretching motion calculated with the NEO-HF, NEO-DFT, and FGH methods for the $[He\text{-}X\text{-}He]^+$ systems with X = H, D, or T. The cc-pVDZ or cc-pVTZ electronic basis set was used as indicated. The NEO-HF and NEO-DFT calculations were performed using a single 1*s* nuclear basis set with a variationally optimized exponent, and the NEO-DFT calculations were performed using two geminals with parameters $(b_1, \gamma_1) = (0.2826, 0.34)$ and $(b_2, \gamma_2) = (0.6346, 2.47)$ for the electron-proton functional derived previously.[1]

References

(1) A. Chakraborty, M. V. Pak, and S. Hammes-Schiffer, *Physical Review Letters* **101**, 153001 (2008).

Appendix D

Supporting Information for Chapter 5[‡]

D.1. Summary of Contents

In this Appendix, we include supporting information corresponding to the material presented in Chapter 5. We include a derivation of the electron-proton functional discussed in Section 5.2B using the adiabatic connection formula.

D.2. Derivation of electron-proton functional using adiabatic connection formula

We consider a scaled Hamiltonian written in terms of a coupling strength, $\lambda > 0$:

$$H = T + \lambda V \tag{D.1}$$

where T and V are given by Eqs. (5.5) and (5.6), respectively. The associated universal functional is given by

$$F_\lambda \left[\rho^e, \rho^p \right] = \min_{\Psi \to \rho^e \rho^p} \left\langle \Psi \middle| T + \lambda V \middle| \Psi \right\rangle = \left\langle \Psi^{\rho^e \rho^p}_{\min,\lambda} \middle| T + \lambda V \middle| \Psi^{\rho^e \rho^p}_{\min,\lambda} \right\rangle, \tag{D.2}$$

[‡]Reproduced with permission from:
A. Sirjoosingh, M. V. Pak, and S. Hammes-Schiffer, "Multicomponent density functional theory study of the interplay between electron-electron and electron-proton correlation," *J. Chem. Phys.* **136**, 174114 (2012). © 2012 American Institute of Physics

where the minimizing wavefunction is denoted $\Psi^{\rho^e\rho^p}_{\min,\lambda}$. The noninteracting system corresponds to $\lambda = 0$ (i.e., $F_0\left[\rho^e,\rho^p\right] = T_s\left[\rho^e,\rho^p\right]$), and the universal functional for the system of interest ($\lambda = 1$) can be expressed by rewriting Eq. (5.8) as

$$\begin{aligned} F_1\left[\rho^e,\rho^p\right] = F_0\left[\rho^e,\rho^p\right] + J_{ee}\left[\rho^e\right] + E_{exc}\left[\rho^e\right] + J_{pp}\left[\rho^p\right] + E_{pxc}\left[\rho^p\right] \\ + J_{ep}\left[\rho^e,\rho^p\right] + E_{epc}\left[\rho^e,\rho^p\right]. \end{aligned} \tag{D.3}$$

Assuming that F_λ is a continuous and differentiable function of λ, this equation can be rewritten as

$$E_{epc}\left[\rho^e,\rho^p\right] = \int_0^1 d\lambda \frac{\partial F_\lambda}{\partial \lambda} - J_{ep}\left[\rho^e,\rho^p\right] - J_{ee}\left[\rho^e\right] - E_{exc}\left[\rho^e\right] - J_{pp}\left[\rho^p\right] - E_{pxc}\left[\rho^p\right]. \tag{D.4}$$

Taking the derivative with respect to λ and using the Hellmann-Feynman theorem leads to

$$\begin{aligned} E_{epc}\left[\rho^e,\rho^p\right] = \int_0^1 d\lambda \left\langle \Psi^{\rho^e\rho^p}_{\min,\lambda} \middle| V \middle| \Psi^{\rho^e\rho^p}_{\min,\lambda} \right\rangle - J_{ep}\left[\rho^e,\rho^p\right] \\ - J_{ee}\left[\rho^e\right] - E_{exc}\left[\rho^e\right] - J_{pp}\left[\rho^p\right] - E_{pxc}\left[\rho^p\right]. \end{aligned} \tag{D.5}$$

The analogous adiabatic connection formula for electronic DFT is[1]

$$E_{exc}\left[\rho^e\right] = \int_0^1 d\lambda \left\langle \Psi^{\rho^e}_{\min,\lambda} \middle| V_{ee} \middle| \Psi^{\rho^e}_{\min,\lambda} \right\rangle - J_{ee}\left[\rho^e\right], \tag{D.6}$$

where $\Psi^{\rho^e}_{\min,\lambda}$ is defined by

$$\left\langle \Psi^{\rho^e}_{\min,\lambda} \middle| T_{ee} + \lambda V_{ee} \middle| \Psi^{\rho^e}_{\min,\lambda} \right\rangle = \min_{\Psi\to\rho^e} \left\langle \Psi \middle| T_{ee} + \lambda V_{ee} \middle| \Psi \right\rangle. \tag{D.7}$$

In this case, the wavefunctions depend only on the electron coordinates. An analogous expression for $E_{pxc}\left[\rho^p\right]$ is valid for the quantum protons.

Substituting Eq. (5.6), as well as Eq. (D.6) and its analog for protons, into Eq. (D.5) leads to

$$\begin{aligned} E_{epc}\left[\rho^e,\rho^p\right] &= \int_0^1 d\lambda \left\langle \Psi^{\rho^e\rho^p}_{\min,\lambda} \middle| V_{ep} \middle| \Psi^{\rho^e\rho^p}_{\min,\lambda} \right\rangle - J_{ep}\left[\rho^e,\rho^p\right] \\ &+ \left(\int_0^1 d\lambda \left\langle \Psi^{\rho^e\rho^p}_{\min,\lambda} \middle| V_{ee} \middle| \Psi^{\rho^e\rho^p}_{\min,\lambda} \right\rangle - \int_0^1 d\lambda \left\langle \Psi^{\rho^e}_{\min,\lambda} \middle| V_{ee} \middle| \Psi^{\rho^e}_{\min,\lambda} \right\rangle \right) \\ &+ \left(\int_0^1 d\lambda \left\langle \Psi^{\rho^e\rho^p}_{\min,\lambda} \middle| V_{pp} \middle| \Psi^{\rho^e\rho^p}_{\min,\lambda} \right\rangle - \int_0^1 d\lambda \left\langle \Psi^{\rho^p}_{\min,\lambda} \middle| V_{pp} \middle| \Psi^{\rho^p}_{\min,\lambda} \right\rangle \right). \end{aligned} \tag{D.8}$$

If we neglect indirect contributions to $E_{epc}\left[\rho^e,\rho^p\right]$ from electron-electron and proton-proton interactions, then the last two lines in Eq. (D.8) vanish, and the final expression for the electron-proton correlation functional is given by

$$E_{epc}\left[\rho^e,\rho^p\right] = \int_0^1 d\lambda \left\langle \Psi^{\rho^e\rho^p}_{\min,\lambda} \middle| V_{ep} \middle| \Psi^{\rho^e\rho^p}_{\min,\lambda} \right\rangle - J_{ep}\left[\rho^e,\rho^p\right]. \tag{D.9}$$

As in Eq. (5.16), we express the quantity in the integrand at any value of λ in terms of an explicitly correlated electron-proton pair density obtained from the minimizing wavefunction in Eq. (D.9), leading to the general expression

$$E_{epc}\left[\rho^e,\rho^p\right] = -\int_0^1 d\lambda \iint \mathrm{d}\mathbf{r}_1^e \mathrm{d}\mathbf{r}_1^p \, \frac{\rho_\lambda^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)}{\left|\mathbf{r}_1^e-\mathbf{r}_1^p\right|} - J_{ep}\left[\rho^e,\rho^p\right]. \tag{D.10}$$

Note that Eq. (D.10) includes $T_{epc}\left[\rho^e,\rho^p\right]$, the kinetic energy contribution due to electron-proton correlation. An expression for the kinetic energy part of the functional is obtained by substituting Eqs. (5.16) and (D.10) into the relation $T_{epc}\left[\rho^e,\rho^p\right] = E_{epc}\left[\rho^e,\rho^p\right] - V_{epc}\left[\rho^e,\rho^p\right]$:

$$T_{epc}\left[\rho^e,\rho^p\right] = -\int_0^1 d\lambda \iint \mathrm{d}\mathbf{r}_1^e \mathrm{d}\mathbf{r}_1^p \, \frac{\rho_\lambda^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)}{\left|\mathbf{r}_1^e-\mathbf{r}_1^p\right|} + \iint \mathrm{d}\mathbf{r}_1^e \mathrm{d}\mathbf{r}_1^p \, \frac{\rho^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)}{\left|\mathbf{r}_1^e-\mathbf{r}_1^p\right|}. \tag{D.11}$$

We now construct a specific electron-proton correlation functional that includes kinetic energy contributions from electron-proton interactions. We denote this functional EPC2-KE because it is derived from the NEO-XCHF2 ansatz for an explicitly correlated wavefunction[2] and includes the kinetic energy contribution. To evaluate Eq. (D.9), we assume that the minimizing wavefunction at each value of λ is given by a NEO-XCHF2-type wavefunction:

$$\Psi_{\min,\lambda}^{\rho^e\rho^p}\left(\mathbf{x}^e,\mathbf{x}^p\right)=\Phi^e\left(\mathbf{x}^e\right)\Phi^p\left(\mathbf{x}^p\right)\sqrt{1+\lambda G\left(\mathbf{r}^e,\mathbf{r}^p\right)}, \tag{D.12}$$

where G is given by Eqs. (5.20) and (5.21) and is fixed by a single set of $\{b_k,\gamma_k\}$ parameters that are independent of λ. The electron-proton pair density of this wavefunction at any value of λ is defined as

$$\rho_\lambda^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)=\frac{N_eN_p}{\left\langle\Psi_{\min,\lambda}^{\rho^e\rho^p}\middle|\Psi_{\min,\lambda}^{\rho^e\rho^p}\right\rangle}\left\langle\Psi_{\min,\lambda}^{\rho^e\rho^p}\middle|\Psi_{\min,\lambda}^{\rho^e\rho^p}\right\rangle_{-e_1p_1} \tag{D.13}$$

and can be evaluated exactly as it was in Ref. 2 for the NEO-XCHF2 wavefunction ($\lambda=1$). The resulting approximate electron-proton pair density expression at any value of λ is given by

$$\begin{aligned}\rho_\lambda^{ep}\left(\mathbf{r}_1^e,\mathbf{r}_1^p\right)\approx\rho^e\rho^p\Bigg[&1+\lambda\left\langle N_e^{-1}N_p^{-1}\rho^e\rho^pg\right\rangle_{ep}-\lambda\left\langle N_e^{-1}\rho^eg\right\rangle_e-\lambda\left\langle N_p^{-1}\rho^pg\right\rangle_p\\&+\frac{\lambda g+\lambda^2\left\langle\rho^eg\right\rangle_e\left\langle\rho^pg\right\rangle_p}{1+\lambda\left\langle\rho^e\rho^pg\right\rangle_{ep}}\Bigg].\end{aligned} \tag{D.14}$$

Substituting this expression into Eq. (D.10), reversing the order of integration, and integrating over λ provides the final expression for the EPC2-KE functional given in Eqs. (5.28) and (5.29).

References

(1) R. G. Parr and W. T. Yang, *Density-Functional Theory of Atoms and Molecules*. (Oxford Science, New York, 1989).
(2) C. Ko, M. V. Pak, C. Swalina, and S. Hammes-Schiffer, *Journal of Chemical Physics* **135**, 054106 (2011).

Appendix E

Supporting Information for Chapter 6[‡]

E.1. Summary of Contents

In this Appendix, we include supporting information corresponding to the material presented in Chapter 6. We include the Fock expressions in the spin orbital, spatial orbital, and atomic orbital bases for RXCHF-fe, RXCHF-ae and RXCHF-ne. We also include a derivation of the one-electron RXCHF-fe energy, a description of an alternative orthogonalization scheme, and a discussion of the operator symmetrization procedure.

E.2. RXCHF-fe Fock Expressions

In this section, we report expressions for the Fock operators in the spin-orbital, spatial-orbital, and atomic-orbital bases for the RXCHF-fe method. The operators defined in these sections are all expressed in their symmetrized forms denoted by a tilde. Details of the symmetrization procedure for all of the RXCHF methods are given in Section E.7.

[‡]Reproduced with permission from:
A. Sirjoosingh, M. V. Pak, C. Swalina, and S. Hammes-Schiffer, "Reduced explicitly correlated Hartree-Fock approach within the nuclear-electronic orbital framework: Theoretical formulation," *J. Chem. Phys.* **139**, 034102 (2013). © 2013 American Institute of Physics

1. Spin-orbital basis

The quantum nuclear Fock operator is obtained by varying the energy in Eq. (6.23) with respect to χ^p, leading to

$$\begin{aligned} f^p\left(\mathbf{x}^p\right) = \frac{1}{S^{\text{RXCHF-fe}}}\Bigg(& \left\langle \chi_1^e \middle| \tilde{\Omega}_1 \middle| \chi_1^e \right\rangle + \sum_{a=2}^{N}\left[\left\langle \chi_1^e \chi_a^e \middle| \tilde{\Omega}_2^{(1)} \middle| \chi_1^e \chi_a^e \right\rangle - \left\langle \chi_1^e \chi_a^e \middle| \tilde{\Omega}_2^{(2)} \middle| \chi_a^e \chi_1^e \right\rangle\right] \\ & + \sum_{a=2}^{N}\sum_{b=2}^{N}\Bigg[\frac{1}{2}\left\langle \chi_1^e \chi_a^e \chi_b^e \middle| \tilde{\Omega}_3^{(1)} \middle| \chi_1^e \chi_a^e \chi_b^e \right\rangle - \frac{1}{2}\left\langle \chi_1^e \chi_a^e \chi_b^e \middle| \tilde{\Omega}_3^{(1)} \middle| \chi_1^e \chi_b^e \chi_a^e \right\rangle \\ & - \left\langle \chi_1^e \chi_a^e \chi_b^e \middle| \tilde{\Omega}_3^{(2)} \middle| \chi_a^e \chi_1^e \chi_b^e \right\rangle + \left\langle \chi_1^e \chi_a^e \chi_b^e \middle| \tilde{\Omega}_3^{(2)} \middle| \chi_b^e \chi_1^e \chi_a^e \right\rangle\Bigg]\Bigg) \\ & - \frac{\left\langle \Psi^{\text{RXCHF-fe}} \middle| H \middle| \Psi^{\text{RXCHF-fe}} \right\rangle}{\left(S^{\text{RXCHF-fe}}\right)^2}\left(\left\langle \chi_1^e \middle| g^2 \middle| \chi_1^e \right\rangle - \sum_{a=2}^{N}\left\langle \chi_1^e \chi_a^e \middle| g(1,p)\, g(2,p) \middle| \chi_a^e \chi_1^e \right\rangle\right). \end{aligned} \tag{E.1}$$

In this expression and those that follow, when explicit dependence of the $\tilde{\Omega}_i^{(j)}$ operators is omitted, we assume that the dependent variables are ordered as $p,1,2,\ldots$, e.g., $\tilde{\Omega}_2^{(1)} \equiv \tilde{\Omega}_2^{(1)}(p,1,2)$.

The special electronic Fock operator is obtained by varying the energy in Eq. (6.23) with respect to χ_1^e, leading to

$$\begin{aligned} f^{e_1}\left(\mathbf{x}_1^e\right) = \frac{1}{S^{\text{RXCHF-fe}}}\Bigg(& \left\langle \chi^p \middle| \tilde{\Omega}_1 \middle| \chi^p \right\rangle + \sum_{a=2}^{N}\left[\left\langle \chi^p \chi_a^e \middle| \tilde{\Omega}_2^{(1)} \middle| \chi^p \chi_a^e \right\rangle - \left\langle \chi^p \chi_a^e \middle| \tilde{\Omega}_2^{(2)} P_2(1,2) \middle| \chi^p \chi_a^e \right\rangle\right] \\ & + \sum_{a=2}^{N}\sum_{b=2}^{N}\Bigg[\frac{1}{2}\left\langle \chi^p \chi_a^e \chi_b^e \middle| \tilde{\Omega}_3^{(1)} \middle| \chi^p \chi_a^e \chi_b^e \right\rangle - \frac{1}{2}\left\langle \chi^p \chi_a^e \chi_b^e \middle| \tilde{\Omega}_3^{(1)} \middle| \chi^p \chi_b^e \chi_a^e \right\rangle \\ & - \left\langle \chi^p \chi_a^e \chi_b^e \middle| \tilde{\Omega}_3^{(2)} P_2(1,2) \middle| \chi^p \chi_a^e \chi_b^e \right\rangle + \left\langle \chi^p \chi_a^e \chi_b^e \middle| \tilde{\Omega}_3^{(2)} P_2(1,2) \middle| \chi^p \chi_b^e \chi_a^e \right\rangle\Bigg]\Bigg) \\ & - \frac{\left\langle \Psi^{\text{RXCHF-fe}} \middle| H \middle| \Psi^{\text{RXCHF-fe}} \right\rangle}{\left(S^{\text{RXCHF-fe}}\right)^2}\Bigg(\left\langle \chi^p \middle| g^2 \middle| \chi^p \right\rangle \\ & - \sum_{a=2}^{N}\left\langle \chi^p \chi_a^e \middle| g(1,p)\, g(2,p)\, P_2(1,2) \middle| \chi^p \chi_a^e \right\rangle\Bigg). \end{aligned} \tag{E.2}$$

In this expression and those that follow, $P_2(i,j)$ is the second-order permutation that exchanges electronic indices i and j.

The regular electronic Fock operator is obtained by varying the energy in Eq. (6.23) with respect to some χ_μ^e, $\mu \in \{2,\ldots N\}$. A compact expression for f^e analogous to Eqs. (E.1) and (E.2) is not easily obtainable because the procedures for operator symmetrizations used in the NEO-XCHF method[1] are not applicable to the fully antisymmetric RXCHF-fe ansatz. A discussion of the operator symmetrizations used for some of the quantities in this paper is included Section E.7. Assuming real spin orbitals, we can express f^e in the spin orbital basis as

$$\begin{aligned}
\left\langle \chi_\mu^e \middle| f^e\left(\mathbf{x}_2^e\right) \middle| \chi_\nu^e \right\rangle = \frac{1}{S^{\text{RXCHF-fe}}} \Bigg(& \left\langle \chi^p \chi_1^e \chi_\mu^e \middle| \tilde{\Omega}_2^{(1)} \middle| \chi^p \chi_1^e \chi_\nu^e \right\rangle - \left\langle \chi^p \chi_1^e \chi_\mu^e \middle| \tilde{\Omega}_2^{(2)} \middle| \chi^p \chi_\nu^e \chi_1^e \right\rangle \\
& + \frac{1}{2} \sum_{a=2}^{N} \Big[2 \left\langle \chi^p \chi_1^e \chi_\mu^e \chi_a^e \middle| \tilde{\Omega}_3^{(1)} \middle| \chi^p \chi_1^e \chi_\nu^e \chi_a^e \right\rangle - 2 \left\langle \chi^p \chi_1^e \chi_\mu^e \chi_a^e \middle| \tilde{\Omega}_3^{(1)} \middle| \chi^p \chi_1^e \chi_a^e \chi_\nu^e \right\rangle \\
& - \left\langle \chi^p \chi_1^e \chi_\mu^e \chi_a^e \middle| \tilde{\Omega}_3^{(2)} \middle| \chi^p \chi_\nu^e \chi_1^e \chi_a^e \right\rangle + \left\langle \chi^p \chi_1^e \chi_\mu^e \chi_a^e \middle| \tilde{\Omega}_3^{(2)} \middle| \chi^p \chi_a^e \chi_1^e \chi_\nu^e \right\rangle \\
& - \left\langle \chi^p \chi_1^e \chi_\nu^e \chi_a^e \middle| \tilde{\Omega}_3^{(2)} \middle| \chi^p \chi_\mu^e \chi_1^e \chi_a^e \right\rangle + \left\langle \chi^p \chi_1^e \chi_\nu^e \chi_a^e \middle| \tilde{\Omega}_3^{(2)} \middle| \chi^p \chi_a^e \chi_1^e \chi_\mu^e \right\rangle \\
& - \left\langle \chi^p \chi_1^e \chi_a^e \chi_\mu^e \middle| \tilde{\Omega}_3^{(2)} \middle| \chi^p \chi_a^e \chi_1^e \chi_\nu^e \right\rangle + \left\langle \chi^p \chi_1^e \chi_a^e \chi_\mu^e \middle| \tilde{\Omega}_3^{(2)} \middle| \chi^p \chi_\nu^e \chi_1^e \chi_a^e \right\rangle \\
& - \left\langle \chi^p \chi_1^e \chi_a^e \chi_\nu^e \middle| \tilde{\Omega}_3^{(2)} \middle| \chi^p \chi_a^e \chi_1^e \chi_\mu^e \right\rangle + \left\langle \chi^p \chi_1^e \chi_a^e \chi_\nu^e \middle| \tilde{\Omega}_3^{(2)} \middle| \chi^p \chi_\mu^e \chi_1^e \chi_a^e \right\rangle \Big] \Bigg) \\
& + \frac{\left\langle \Psi^{\text{RXCHF-fe}} \middle| H \middle| \Psi^{\text{RXCHF-fe}} \right\rangle}{\left(S^{\text{RXCHF-fe}}\right)^2} \left\langle \chi^p \chi_1^e \chi_\mu^e \middle| g(1,p)\, g(2,p) \middle| \chi^p \chi_\nu^e \chi_1^e \right\rangle
\end{aligned} \tag{E.3}$$

for any spin orbitals χ_μ^e, χ_ν^e with $\mu,\nu \in \{2,\ldots N\}$.

2. Spatial-orbital basis

Next we integrate the RXCHF-fe Fock operators in Eqs. (E.1) - (E.3) over spin to obtain the spatial Fock operators. The quantum nuclear Fock operator is given by

$$
\begin{aligned}
f^{p}\left(\mathbf{r}^{p}\right) = \frac{1}{S^{\text{RXCHF-fe}}}\Bigg(& \left\langle \psi_1^e \middle| \tilde{\Omega}_1 \middle| \psi_1^e \right\rangle + \sum_a^{N_\alpha}\left[\left\langle \psi_1^e \psi_a^\alpha \middle| \tilde{\Omega}_2^{(1)} \middle| \psi_1^e \psi_a^a \right\rangle - \left\langle \psi_1^e \psi_a^\alpha \middle| \tilde{\Omega}_2^{(2)} \middle| \psi_a^\alpha \psi_1^e \right\rangle \right] \\
& + \sum_a^{N_\beta} \left\langle \psi_1^e \psi_a^\beta \middle| \tilde{\Omega}_2^{(1)} \middle| \psi_1^e \psi_a^\beta \right\rangle \\
& + \sum_a^{N_\alpha}\sum_b^{N_\alpha}\Bigg[\frac{1}{2}\left\langle \psi_1^e \psi_a^\alpha \psi_b^\alpha \middle| \tilde{\Omega}_3^{(1)} \middle| \psi_1^e \psi_a^a \psi_b^a \right\rangle - \frac{1}{2}\left\langle \psi_1^e \psi_a^a \psi_b^a \middle| \tilde{\Omega}_3^{(1)} \middle| \psi_1^e \psi_b^\alpha \psi_a^\alpha \right\rangle \\
& - \left\langle \psi_1^e \psi_a^\alpha \psi_b^\alpha \middle| \tilde{\Omega}_3^{(2)} \middle| \psi_a^a \psi_1^e \psi_b^\alpha \right\rangle + \left\langle \psi_1^e \psi_a^\alpha \psi_b^\alpha \middle| \tilde{\Omega}_3^{(2)} \middle| \psi_b^\alpha \psi_1^e \psi_a^\alpha \right\rangle \Bigg] \\
& + \sum_a^{N_\alpha}\sum_b^{N_\beta}\Bigg[\frac{1}{2}\left\langle \psi_1^e \psi_a^\alpha \psi_b^\beta \middle| \tilde{\Omega}_3^{(1)} \middle| \psi_1^e \psi_a^a \psi_b^\beta \right\rangle - \left\langle \psi_1^e \psi_a^\alpha \psi_b^\beta \middle| \tilde{\Omega}_3^{(2)} \middle| \psi_a^a \psi_1^e \psi_b^\beta \right\rangle \Bigg] \\
& + \sum_a^{N_\beta}\sum_b^{N_\alpha} \frac{1}{2}\left\langle \psi_1^e \psi_a^\beta \psi_b^\alpha \middle| \tilde{\Omega}_3^{(1)} \middle| \psi_1^e \psi_a^\beta \psi_b^a \right\rangle \\
& + \sum_a^{N_\alpha}\sum_b^{N_\alpha}\Bigg[\frac{1}{2}\left\langle \psi_1^e \psi_a^\beta \psi_b^\beta \middle| \tilde{\Omega}_3^{(1)} \middle| \psi_1^e \psi_a^\beta \psi_b^\beta \right\rangle - \frac{1}{2}\left\langle \psi_1^e \psi_a^\beta \psi_b^\beta \middle| \tilde{\Omega}_3^{(1)} \middle| \psi_1^e \psi_b^\beta \psi_a^\beta \right\rangle \Bigg] \Bigg) \\
& - \frac{\left\langle \Psi^{\text{RXCHF-fe}} \middle| H \middle| \Psi^{\text{RXCHF-fe}} \right\rangle}{\left(S^{\text{RXCHF-fe}}\right)^2} \Bigg(\left\langle \psi_1^e \middle| g^2 \middle| \psi_1^e \right\rangle - \sum_a^{N_\alpha} \left\langle \psi_1^e \psi_a^\alpha \middle| g(1,p)\, g(2,p) \middle| \psi_a^\alpha \psi_1^e \right\rangle \Bigg). \qquad \text{(E.4)}
\end{aligned}
$$

The special electronic Fock operator is given by

$$
\begin{aligned}
f^{e_1}\left(\mathbf{r}_1^e\right) = \frac{1}{S^{\text{RXCHF-fe}}}\Bigg(& \left\langle \psi^p \middle| \tilde{\Omega}_1 \middle| \psi^p \right\rangle + \sum_a^{N_\alpha}\left[\left\langle \psi^p \psi_a^\alpha \middle| \tilde{\Omega}_2^{(1)} \middle| \psi^p \psi_a^\alpha \right\rangle - \left\langle \psi^p \psi_a^\alpha \middle| \tilde{\Omega}_2^{(2)} P_2(1,2) \middle| \psi^p \psi_a^\alpha \right\rangle \right] \\
& + \sum_a^{N_\beta} \left\langle \psi^p \psi_a^\beta \middle| \tilde{\Omega}_2^{(1)} \middle| \psi^p \psi_a^\beta \right\rangle \\
& + \sum_a^{N_\alpha}\sum_b^{N_\alpha}\left[\frac{1}{2}\left\langle \psi^p \psi_a^\alpha \psi_b^\alpha \middle| \tilde{\Omega}_3^{(1)} \middle| \psi^p \psi_a^\alpha \psi_b^\alpha \right\rangle - \frac{1}{2}\left\langle \psi^p \psi_a^\alpha \psi_b^\alpha \middle| \tilde{\Omega}_3^{(1)} \middle| \psi^p \psi_b^\alpha \psi_a^\alpha \right\rangle \right. \\
& \left. - \left\langle \psi^p \psi_a^\alpha \psi_b^\alpha \middle| \tilde{\Omega}_3^{(2)} P_2(1,2) \middle| \psi^p \psi_a^\alpha \psi_b^\alpha \right\rangle + \left\langle \psi^p \psi_a^\alpha \psi_b^\alpha \middle| \tilde{\Omega}_3^{(2)} P_2(1,2) \middle| \psi^p \psi_b^\alpha \psi_a^\alpha \right\rangle \right] \\
& + \sum_a^{N_\alpha}\sum_b^{N_\beta}\left[\frac{1}{2}\left\langle \psi^p \psi_a^\alpha \psi_b^\beta \middle| \tilde{\Omega}_3^{(1)} \middle| \psi^p \psi_a^\alpha \psi_b^\beta \right\rangle - \left\langle \psi^p \psi_a^\alpha \psi_b^\beta \middle| \tilde{\Omega}_3^{(2)} P_2(1,2) \middle| \psi^p \psi_a^\alpha \psi_b^\beta \right\rangle \right] \\
& + \sum_a^{N_\beta}\sum_b^{N_\alpha} \frac{1}{2}\left\langle \psi^p \psi_a^\beta \psi_b^\alpha \middle| \tilde{\Omega}_3^{(1)} \middle| \psi^p \psi_a^\beta \psi_b^\alpha \right\rangle \\
& + \sum_a^{N_\beta}\sum_b^{N_\beta}\left[\frac{1}{2}\left\langle \psi^p \psi_a^\beta \psi_b^\beta \middle| \tilde{\Omega}_3^{(1)} \middle| \psi^p \psi_a^\beta \psi_b^\beta \right\rangle - \frac{1}{2}\left\langle \psi^p \psi_a^\beta \psi_b^\beta \middle| \tilde{\Omega}_3^{(1)} \middle| \psi^p \psi_b^\beta \psi_a^\beta \right\rangle \right] \Bigg) \\
& - \frac{\left\langle \Psi^{\text{RXCHF-fe}} \middle| H \middle| \Psi^{\text{RXCHF-fe}} \right\rangle}{\left(S^{\text{RXCHF-fe}}\right)^2}\Bigg(\left\langle \psi^p \middle| g^2 \middle| \psi^p \right\rangle \\
& - \sum_a^{N_\alpha} \left\langle \psi^p \psi_a^\alpha \middle| g(1,p)\, g(2,p)\, P_2(1,2) \middle| \psi^p \psi_a^\alpha \right\rangle \Bigg).
\end{aligned}
\tag{E.5}
$$

Integrating Eq. (E.3) over spin results in different expressions for α and β spins. This situation arises for the fully antisymmetric RXCHF-fe ansatz because regular electrons in α-spin orbitals interact differently with the geminal-coupled electronic orbital than do regular electrons in β-spin orbitals. The regular electronic Fock operator for α-spin orbitals, f^α, in the spatial orbital basis is

$$
\begin{aligned}
\left\langle \psi_\mu^\alpha \middle| f^\alpha\left(\mathbf{r}_2^e\right) \middle| \psi_\nu^\alpha \right\rangle &= \int \mathrm{d}\omega_2^e \psi_\mu^\alpha\left(\mathbf{r}_2^e\right)\alpha\left(\omega_2^e\right) f^e\left(\mathbf{r}_2^e,\omega_2^e\right)\psi_\nu^\alpha\left(\mathbf{r}_2^e\right)\alpha\left(\omega_2^e\right) \\
&= \frac{1}{S^{\text{RXCHF-fe}}}\Bigg(\left\langle \psi^p\psi_1^e\psi_\mu^\alpha \middle| \tilde{\Omega}_2^{(1)} \middle| \psi^p\psi_1^e\psi_\nu^\alpha \right\rangle - \left\langle \psi^p\psi_1^e\psi_\mu^\alpha \middle| \tilde{\Omega}_2^{(2)} \middle| \psi^p\psi_\nu^\alpha\psi_1^e \right\rangle \\
&+ \frac{1}{2}\sum_a^{N_\alpha} \\
&\Bigg[2\left\langle \psi^p\psi_1^e\psi_\mu^\alpha\psi_a^\alpha \middle| \tilde{\Omega}_3^{(1)} \middle| \psi^p\psi_1^e\psi_\nu^\alpha\psi_a^\alpha \right\rangle - 2\left\langle \psi^p\psi_1^e\psi_\mu^\alpha\psi_a^\alpha \middle| \tilde{\Omega}_3^{(1)} \middle| \psi^p\psi_1^e\psi_a^\alpha\psi_\nu^\alpha \right\rangle \\
&- \left\langle \psi^p\psi_1^e\psi_\mu^\alpha\psi_a^\alpha \middle| \tilde{\Omega}_3^{(2)} \middle| \psi^p\psi_\nu^\alpha\psi_1^e\psi_a^\alpha \right\rangle + \left\langle \psi^p\psi_1^e\psi_\mu^\alpha\psi_a^\alpha \middle| \tilde{\Omega}_3^{(2)} \middle| \psi^p\psi_a^\alpha\psi_1^e\psi_\nu^\alpha \right\rangle \\
&- \left\langle \psi^p\psi_1^e\psi_\nu^\alpha\psi_a^\alpha \middle| \tilde{\Omega}_3^{(2)} \middle| \psi^p\psi_\mu^\alpha\psi_1^e\psi_a^\alpha \right\rangle + \left\langle \psi^p\psi_1^e\psi_\nu^\alpha\psi_a^\alpha \middle| \tilde{\Omega}_3^{(2)} \middle| \psi^p\psi_a^\alpha\psi_1^e\psi_\mu^\alpha \right\rangle \\
&- \left\langle \psi^p\psi_1^e\psi_a^\alpha\psi_\mu^\alpha \middle| \tilde{\Omega}_3^{(2)} \middle| \psi^p\psi_a^\alpha\psi_1^e\psi_\nu^\alpha \right\rangle + \left\langle \psi^p\psi_1^e\psi_a^\alpha\psi_\mu^\alpha \middle| \tilde{\Omega}_3^{(2)} \middle| \psi^p\psi_\nu^\alpha\psi_1^e\psi_a^\alpha \right\rangle \\
&- \left\langle \psi^p\psi_1^e\psi_a^\alpha\psi_\nu^\alpha \middle| \tilde{\Omega}_3^{(2)} \middle| \psi^p\psi_a^\alpha\psi_1^e\psi_\mu^\alpha \right\rangle + \left\langle \psi^p\psi_1^e\psi_a^\alpha\psi_\nu^\alpha \middle| \tilde{\Omega}_3^{(2)} \middle| \psi^p\psi_\mu^\alpha\psi_1^e\psi_a^\alpha \right\rangle \Bigg] \\
&+ \frac{1}{2}\sum_a^{N_\beta} \Bigg[2\left\langle \psi^p\psi_1^e\psi_\mu^\alpha\psi_a^\beta \middle| \tilde{\Omega}_3^{(1)} \middle| \psi^p\psi_1^e\psi_\nu^\alpha\psi_a^\beta \right\rangle \\
&- \left\langle \psi^p\psi_1^e\psi_\mu^\alpha\psi_a^\beta \middle| \tilde{\Omega}_3^{(2)} \middle| \psi^p\psi_\nu^\alpha\psi_1^e\psi_a^\beta \right\rangle - \left\langle \psi^p\psi_1^e\psi_\nu^\alpha\psi_a^\beta \middle| \tilde{\Omega}_3^{(2)} \middle| \psi^p\psi_\mu^\alpha\psi_1^e\psi_a^\beta \right\rangle \Bigg] \Bigg) \\
&+ \frac{\left\langle \Psi^{\text{RXCHF-fe}} \middle| H \middle| \Psi^{\text{RXCHF-fe}} \right\rangle}{\left(S^{\text{RXCHF-fe}}\right)^2} \left\langle \psi^p\psi_1^e\psi_\mu^\alpha \middle| g\left(1,p\right) g\left(2,p\right) \middle| \psi^p\psi_\nu^\alpha\psi_1^e \right\rangle .
\end{aligned}
\tag{E.6}
$$

The regular electronic Fock operator for β-spin orbitals, f^β, in the spatial orbital basis is

$$
\begin{aligned}
\left\langle \psi_\mu^\beta \middle| f^\beta\left(\mathbf{r}_2^e\right) \middle| \psi_\nu^\beta \right\rangle &= \int \mathrm{d}\omega_2^e \psi_\mu^\beta\left(\mathbf{r}_2^e\right)\beta\left(\omega_2^e\right) f^e\left(\mathbf{r}_2^e,\omega_2^e\right)\psi_\nu^\beta\left(\mathbf{r}_2^e\right)\beta\left(\omega_2^e\right) \\
&= \frac{1}{S^{\text{RXCHF-fe}}}\Bigg(\left\langle \psi^p\psi_1^e\psi_\mu^\beta \middle| \tilde{\Omega}_2^{(1)} \middle| \psi^p\psi_1^e\psi_\nu^\beta \right\rangle \\
&+ \frac{1}{2}\sum_a^{N_\alpha} \Bigg[2\left\langle \psi^p\psi_1^e\psi_\mu^\beta\psi_a^\alpha \middle| \tilde{\Omega}_3^{(1)} \middle| \psi^p\psi_1^e\psi_\nu^\beta\psi_a^\alpha \right\rangle \\
&- \left\langle \psi^p\psi_1^e\psi_a^\alpha\psi_\mu^\beta \middle| \tilde{\Omega}_3^{(2)} \middle| \psi^p\psi_a^\alpha\psi_1^e\psi_\nu^\beta \right\rangle - \left\langle \psi^p\psi_1^e\psi_a^\alpha\psi_\nu^\beta \middle| \tilde{\Omega}_3^{(2)} \middle| \psi^p\psi_a^\alpha\psi_1^e\psi_\mu^\beta \right\rangle \Bigg] \\
&+ \sum_a^{N_\beta} \Bigg[\left\langle \psi^p\psi_1^e\psi_\mu^\beta\psi_a^\beta \middle| \tilde{\Omega}_3^{(1)} \middle| \psi^p\psi_1^e\psi_\nu^\beta\psi_a^\beta \right\rangle - \left\langle \psi^p\psi_1^e\psi_\mu^\beta\psi_a^\beta \middle| \tilde{\Omega}_3^{(1)} \middle| \psi^p\psi_1^e\psi_a^\beta\psi_\nu^\beta \right\rangle \Bigg] \Bigg).
\end{aligned}
\tag{E.7}
$$

3. Atomic-orbital basis

We now expand the spatial orbitals in the AO bases, as in Eqs. (6.40) – (6.43). We first define density matrices as

$$P^{p}_{\mu'\nu'} = C^{p*}_{\mu'} C^{p}_{\nu'}, \tag{E.8}$$

$$P^{e_1}_{\mu"\nu"} = C^{e_1*}_{\mu"} C^{e_1}_{\nu"}, \tag{E.9}$$

$$P^{\alpha}_{\mu\nu} = \sum_{a}^{N_\alpha} C^{\alpha*}_{\mu a} C^{\alpha}_{\nu a}, \tag{E.10}$$

$$P^{\beta}_{\mu\nu} = \sum_{a}^{N_\beta} C^{\beta*}_{\mu a} C^{\beta}_{\nu a}, \tag{E.11}$$

$$P^{\text{tot}}_{\mu\nu} = P^{\alpha}_{\mu\nu} + P^{\beta}_{\mu\nu}. \tag{E.12}$$

The nuclear Fock operator in the AO basis is given by

$$\begin{aligned}
F^{p}_{\mu'\nu'} = \frac{1}{S^{\text{RXCHF-fe}}} \Bigg(& \sum_{\mu"\nu"} P^{e_1}_{\mu"\nu"} \Gamma_1 \left(\mu',\nu';\mu",\nu"\right) + \sum_{\mu\nu}\sum_{\mu"\nu"} P^{e_1}_{\mu"\nu"} \\
& \times \Big[P^{\text{tot}}_{\mu\nu} \Gamma^{(1)}_2 \left(\mu',\nu';\mu",\nu";\mu,\nu\right) - P^{\alpha}_{\mu\nu} \Gamma^{(2)}_2 \left(\mu',\nu';\mu",\nu";\mu,\nu\right) \Big] \\
& + \sum_{\mu_1\nu_1}\sum_{\mu_2\nu_2}\sum_{\mu"\nu"} P^{e_1}_{\mu"\nu"} \\
& \times \Bigg[\frac{1}{2} P^{\text{tot}}_{\mu_1\nu_1} P^{\text{tot}}_{\mu_2\nu_2} \Gamma^{(1a)}_3 \left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right) \\
& - \frac{1}{2} \left(P^{\alpha}_{\mu_1\nu_1} P^{\alpha}_{\mu_2\nu_2} + P^{\beta}_{\mu_1\nu_1} P^{\beta}_{\mu_2\nu_2} \right) \Gamma^{(1b)}_3 \left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right) \\
& - P^{\alpha}_{\mu_1\nu_1} P^{\text{tot}}_{\mu_2\nu_2} \Gamma^{(2a)}_3 \left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right) \\
& + P^{\alpha}_{\mu_1\nu_1} P^{\alpha}_{\mu_2\nu_2} \Gamma^{(2b)}_3 \left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right) \Bigg] \Bigg) \\
& - \frac{\left\langle \Psi^{\text{RXCHF-fe}} \middle| H \middle| \Psi^{\text{RXCHF-fe}} \right\rangle}{\left(S^{\text{RXCHF-fe}}\right)^2} \Bigg(\sum_{\mu"\nu"} P^{e_1}_{\mu"\nu"} \Gamma_{1s} \left(\mu',\nu';\mu",\nu"\right) \\
& - \sum_{\mu\nu}\sum_{\mu"\nu"} P^{e_1}_{\mu"\nu"} P^{\alpha}_{\mu\nu} \Gamma_{2s} \left(\mu',\nu';\mu",\nu";\mu,\nu\right) \Bigg).
\end{aligned} \tag{E.13}$$

The special electronic Fock operator in the AO basis is given by

$$
\begin{aligned}
F^{e_1}_{\mu"\nu"} = \frac{1}{S^{\text{RXCHF-fe}}}\Bigg(& \sum_{\mu'\nu'} P^{p}_{\mu'\nu'} \Gamma_1\left(\mu',\nu';\mu",\nu"\right) + \sum_{\mu\nu}\sum_{\mu'\nu'} P^{p}_{\mu'\nu'} \\
& \times\Big[P^{\text{tot}}_{\mu\nu}\Gamma^{(1)}_2\left(\mu',\nu';\mu",\nu";\mu,\nu\right) - P^{\alpha}_{\mu\nu}\Gamma^{(2)}_2\left(\mu',\nu';\mu",\nu";\mu,\nu\right)\Big] \\
& + \sum_{\mu_1\nu_1}\sum_{\mu_2\nu_2}\sum_{\mu'\nu'} P^{p}_{\mu'\nu'} \\
& \times\Bigg[\frac{1}{2} P^{\text{tot}}_{\mu_1\nu_1} P^{\text{tot}}_{\mu_2\nu_2} \Gamma^{(1a)}_3\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right) \\
& -\frac{1}{2}\left(P^{\alpha}_{\mu_1\nu_1} P^{\alpha}_{\mu_2\nu_2} + P^{\beta}_{\mu_1\nu_1} P^{\beta}_{\mu_2\nu_2}\right)\Gamma^{(1b)}_3\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right) \\
& - P^{\alpha}_{\mu_1\nu_1} P^{\text{tot}}_{\mu_2\nu_2} \Gamma^{(2a)}_3\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right) \\
& + P^{\alpha}_{\mu_1\nu_1} P^{\alpha}_{\mu_2\nu_2} \Gamma^{(2b)}_3\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right)\Bigg]\Bigg) \\
& - \frac{\left\langle \Psi^{\text{RXCHF-fe}} \middle| H \middle| \Psi^{\text{RXCHF-fe}} \right\rangle}{\left(S^{\text{RXCHF-fe}}\right)^2}\Bigg(\sum_{\mu'\nu'} P^{p}_{\mu'\nu'} \Gamma_{1s}\left(\mu',\nu';\mu",\nu"\right) \\
& - \sum_{\mu\nu}\sum_{\mu'\nu'} P^{p}_{\mu'\nu'} P^{\alpha}_{\mu\nu} \Gamma_{2s}\left(\mu',\nu';\mu",\nu";\mu,\nu\right)\Bigg).
\end{aligned}
\tag{E.14}
$$

The regular electronic Fock operator for α-spin orbitals in the AO basis is given by

$$
\begin{aligned}
F^{\alpha}_{\mu_1\nu_1} = \frac{1}{S^{\text{RXCHF-fe}}}\Bigg(& \sum_{\mu'\nu'}\sum_{\mu"\nu"} P^{p}_{\mu'\nu'} P^{e_1}_{\mu"\nu"} \\
& \times\Big[\Gamma^{(1)}_2\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1\right) - \Gamma^{(2)}_2\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1\right)\Big] \\
& + \frac{1}{2}\sum_{\mu_2\nu_2}\sum_{\mu'\nu'}\sum_{\mu"\nu"} P^{p}_{\mu'\nu'} P^{e_1}_{\mu"\nu"} \\
& \times\Big[2P^{\text{tot}}_{\mu_2\nu_2}\Gamma^{(1a)}_3\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right) - 2P^{\alpha}_{\mu_2\nu_2}\Gamma^{(1b)}_3\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right) \\
& - P^{\text{tot}}_{\mu_2\nu_2}\left(\Gamma^{(2a)}_3\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right) + \Gamma^{(2a)}_3\left(\mu',\nu';\mu",\nu";\nu_1,\mu_1;\mu_2,\nu_2\right)\right) \\
& + P^{\alpha}_{\mu_2\nu_2}\left(\Gamma^{(2b)}_3\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right) + \Gamma^{(2b)}_3\left(\mu',\nu';\mu",\nu";\nu_1,\mu_1;\mu_2,\nu_2\right)\right) \\
& - P^{\alpha}_{\mu_2\nu_2}\left(\Gamma^{(2c)}_3\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right) + \Gamma^{(2c)}_3\left(\mu',\nu';\mu",\nu";\nu_1,\mu_1;\mu_2,\nu_2\right)\right) \\
& + P^{\alpha}_{\mu_2\nu_2}\left(\Gamma^{(2d)}_3\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right) + \Gamma^{(2d)}_3\left(\mu',\nu';\mu",\nu";\nu_1,\mu_1;\mu_2,\nu_2\right)\right)\Big]\Bigg) \\
& + \frac{\left\langle \Psi^{\text{RXCHF-fe}} \middle| H \middle| \Psi^{\text{RXCHF-fe}} \right\rangle}{\left(S^{\text{RXCHF-fe}}\right)^2}\sum_{\mu'\nu'}\sum_{\mu"\nu"} P^{p}_{\mu'\nu'} P^{e_1}_{\mu"\nu"} \Gamma_{2s}\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1\right).
\end{aligned}
\tag{E.15}
$$

The regular electronic Fock operator for β-spin orbitals in the AO basis is given by

$$
\begin{aligned}
F^{\beta}_{\mu_1\nu_1} = \frac{1}{S^{\text{RXCHF-fe}}}\Bigg(& \sum_{\mu'\nu'}\sum_{\mu"\nu"} P^{p}_{\mu'\nu'} P^{e_1}_{\mu"\nu"} \Gamma^{(1)}_{2}\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1\right) \\
& + \frac{1}{2}\sum_{\mu_2\nu_2}\sum_{\mu'\nu'}\sum_{\mu"\nu"} P^{p}_{\mu'\nu'} P^{e_1}_{\mu"\nu"} \\
& \times\Big[2P^{\text{tot}}_{\mu_2\nu_2}\Gamma^{(1a)}_{3}\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right) - 2P^{\beta}_{\mu_2\nu_2}\Gamma^{(1b)}_{3}\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right) \\
& - P^{\alpha}_{\mu_2\nu_2}\left(\Gamma^{(2c)}_{3}\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right) + \Gamma^{(2c)}_{3}\left(\mu',\nu';\mu",\nu";\nu_1,\mu_1;\mu_2,\nu_2\right)\right)\Big]\Bigg).
\end{aligned}
\tag{E.16}
$$

The integrals required for the evaluation of these quantities are defined as

$$
\Gamma_{1}\left(\mu',\nu';\mu",\nu"\right) = \left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu"} \middle| \tilde{\Omega}_1 \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu"} \right\rangle, \tag{E.17}
$$

$$
\Gamma_{1s}\left(\mu',\nu';\mu",\nu"\right) = \left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu"} \middle| g^{2} \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu"} \right\rangle, \tag{E.18}
$$

$$
\Gamma^{(1)}_{2}\left(\mu',\nu';\mu",\nu";\mu,\nu\right) = \left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu"}\phi^{e}_{\mu} \middle| \tilde{\Omega}^{(1)}_{2} \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu"}\phi^{e}_{\nu} \right\rangle, \tag{E.19}
$$

$$
\Gamma^{(2)}_{2}\left(\mu',\nu';\mu",\nu";\mu,\nu\right) = \left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu"}\phi^{e}_{\mu} \middle| \tilde{\Omega}^{(2)}_{2} \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu}\phi^{e}_{\nu"} \right\rangle, \tag{E.20}
$$

$$
\Gamma_{2s}\left(\mu',\nu';\mu",\nu";\mu,\nu\right) = \left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu"}\phi^{e}_{\mu} \middle| g(1)g(2) \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu}\phi^{e}_{\nu"} \right\rangle, \tag{E.21}
$$

$$
\Gamma^{(1a)}_{3}\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right) = \left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu"}\phi^{e}_{\mu_1}\phi^{e}_{\mu_2} \middle| \tilde{\Omega}^{(1)}_{3} \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu"}\phi^{e}_{\nu_1}\phi^{e}_{\nu_2} \right\rangle, \tag{E.22}
$$

$$
\Gamma^{(1b)}_{3}\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right) = \left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu"}\phi^{e}_{\mu_1}\phi^{e}_{\mu_2} \middle| \tilde{\Omega}^{(1)}_{3} \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu"}\phi^{e}_{\nu_2}\phi^{e}_{\nu_1} \right\rangle, \tag{E.23}
$$

$$
\Gamma^{(2a)}_{3}\left(\mu',\nu';\mu",\nu";\nu_1,\mu_1;\mu_2,\nu_2\right) = \left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu"}\phi^{e}_{\mu_1}\phi^{e}_{\mu_2} \middle| \tilde{\Omega}^{(2)}_{3} \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu_1}\phi^{e}_{\nu"}\phi^{e}_{\nu_2} \right\rangle, \tag{E.24}
$$

$$
\Gamma^{(2b)}_{3}\left(\mu',\nu';\mu",\nu";\nu_1,\mu_1;\mu_2,\nu_2\right) = \left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu"}\phi^{e}_{\mu_1}\phi^{e}_{\mu_2} \middle| \tilde{\Omega}^{(2)}_{3} \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu_2}\phi^{e}_{\nu"}\phi^{e}_{\nu_1} \right\rangle, \tag{E.25}
$$

$$
\Gamma^{(2c)}_{3}\left(\mu',\nu';\mu",\nu";\nu_1,\mu_1;\mu_2,\nu_2\right) = \left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu"}\phi^{e}_{\mu_2}\phi^{e}_{\mu_1} \middle| \tilde{\Omega}^{(2)}_{3} \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu_2}\phi^{e}_{\nu"}\phi^{e}_{\nu_1} \right\rangle, \tag{E.26}
$$

$$
\Gamma^{(2d)}_{3}\left(\mu',\nu';\mu",\nu";\mu_1,\nu_1;\mu_2,\nu_2\right) = \left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu"}\phi^{e}_{\mu_2}\phi^{e}_{\mu_1} \middle| \tilde{\Omega}^{(2)}_{3} \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu_1}\phi^{e}_{\nu"}\phi^{e}_{\nu_2} \right\rangle. \tag{E.27}
$$

E.3. RXCHF-ne Fock Expressions

In this section, we report expressions for the Fock operators in the spin-orbital, spatial-orbital, and atomic-orbital bases for the RXCHF-ne method.

1. Spin-orbital basis

The quantum nuclear Fock operator is obtained by varying the energy in Eq. (6.24) with respect to χ^p, leading to

$$f^p\left(\mathbf{x}^p\right)=\frac{1}{S^{\text{RXCHF-ne}}}\left(\left\langle\chi_1^e\left|\tilde{\Omega}_1\right|\chi_1^e\right\rangle+\sum_{a=2}^{N}\left\langle\chi_1^e\chi_a^e\left|\tilde{\Omega}_2\right|\chi_1^e\chi_a^e\right\rangle\right)-\frac{E_G}{S^{\text{RXCHF-ne}}}\left\langle\chi_1^e\left|g^2\right|\chi_1^e\right\rangle, \quad \text{(E.28)}$$

where $E_G=E_{G1}+E_{G2}$.

The special electronic Fock operator is obtained by varying the energy in Eq. (6.24) with respect to χ_1^e, leading to

$$f^{e_1}\left(\mathbf{x}_1^e\right)=\frac{1}{S^{\text{RXCHF-ne}}}\left(\left\langle\chi^p\left|\tilde{\Omega}_1\right|\chi^p\right\rangle+\sum_{a=2}^{N}\left\langle\chi^p\chi_a^e\left|\tilde{\Omega}_2\right|\chi^p\chi_a^e\right\rangle\right)-\frac{E_G}{S^{\text{RXCHF-ne}}}\left\langle\chi^p\left|g^2\right|\chi^p\right\rangle. \quad \text{(E.29)}$$

The regular electronic Fock operator is obtained by varying the energy in Eq. (6.24) with respect to some $\chi_\mu^e,\ \ \mu\in\{2,\ldots N\}$, leading to

$$f^e\left(\mathbf{x}_2^e\right)=h^e(2)+\sum_{a=2}^{N}\left(J_a(2)-K_a(2)\right)+\frac{1}{S^{\text{RXCHF-ne}}}\left\langle\chi^p\chi_1^e\left|\tilde{\Omega}_2\right|\chi^p\chi_1^e\right\rangle, \quad \text{(E.30)}$$

where J_a and K_a are the usual electronic Coulomb and exchange operators, respectively.

2. Spatial-orbital basis

Next we integrate the RXCHF-ne Fock operators in Eqs. (E.28) - (E.30) over spin to obtain the spatial Fock operators. Although RXCHF-ne allows the possibility of using a restricted Hartree-Fock (RHF) approach for a closed-shell treatment of regular electrons, we report open-shell expressions using an analogous UHF approach to that used in the RXCHF-fe expressions for generality.

The quantum nuclear Fock operator is given by

$$f^{p}\left(\mathbf{r}^{p}\right)=\frac{1}{S^{\text{RXCHF-ne}}}\left(\left\langle\psi_{1}^{e}\left|\tilde{\Omega}_{1}\right|\psi_{1}^{e}\right\rangle+\sum_{a}^{N_{\alpha}}\left\langle\psi_{1}^{e}\psi_{a}^{\alpha}\left|\tilde{\Omega}_{2}\right|\psi_{1}^{e}\psi_{a}^{a}\right\rangle+\sum_{a}^{N_{\beta}}\left\langle\psi_{1}^{e}\psi_{a}^{\beta}\left|\tilde{\Omega}_{2}\right|\psi_{1}^{e}\psi_{a}^{\beta}\right\rangle\right) \\ -\frac{E_{G}}{S^{\text{RXCHF-ne}}}\left\langle\chi_{1}^{e}\left|g^{2}\right|\chi_{1}^{e}\right\rangle. \tag{E.31}$$

The special electronic Fock operator is given by

$$f^{e_{1}}\left(\mathbf{r}_{1}^{e}\right)=\frac{1}{S^{\text{RXCHF-ne}}}\left(\left\langle\psi^{p}\left|\tilde{\Omega}_{1}\right|\psi^{p}\right\rangle+\sum_{a}^{N_{\alpha}}\left\langle\psi^{p}\psi_{a}^{\alpha}\left|\tilde{\Omega}_{2}\right|\psi^{p}\psi_{a}^{\alpha}\right\rangle+\sum_{a}^{N_{\beta}}\left\langle\psi^{p}\psi_{a}^{\beta}\left|\tilde{\Omega}_{2}\right|\psi^{p}\psi_{a}^{\beta}\right\rangle\right) \\ -\frac{E_{G}}{S^{\text{RXCHF-ne}}}\left\langle\chi^{p}\left|g^{2}\right|\chi^{p}\right\rangle. \tag{E.32}$$

The regular electronic Fock operator for α-spin orbitals is given by

$$f^{\alpha}\left(\mathbf{x}_{2}^{e}\right)=h^{e}\left(2\right)+\sum_{a}^{N_{\alpha}}\left(J_{a}^{\alpha}\left(2\right)-K_{a}^{\alpha}\left(2\right)\right)+\sum_{a}^{N_{\beta}}J_{a}^{\beta}\left(2\right)+\frac{1}{S^{\text{RXCHF-ne}}}\left\langle\psi^{p}\psi_{1}^{e}\left|\tilde{\Omega}_{2}\right|\psi^{p}\psi_{1}^{e}\right\rangle. \tag{E.33}$$

The regular electronic Fock operator for β-spin orbitals is given by

$$f^{\beta}\left(\mathbf{x}_{2}^{e}\right)=h^{e}\left(2\right)+\sum_{a}^{N_{\beta}}\left(J_{a}^{\beta}\left(2\right)-K_{a}^{\beta}\left(2\right)\right)+\sum_{a}^{N_{\alpha}}J_{a}^{\alpha}\left(2\right)+\frac{1}{S^{\text{RXCHF-ne}}}\left\langle\psi^{p}\psi_{1}^{e}\left|\tilde{\Omega}_{2}\right|\psi^{p}\psi_{1}^{e}\right\rangle. \tag{E.34}$$

In the case of $N_{\alpha}=N_{\beta}$, $f^{\alpha}=f^{\beta}$, and it is straightforward to rewrite the modified Fock operators in an RHF formalism.

3. Atomic-orbital basis

We now expand the spatial orbitals in the AO bases, as in Eqs. (6.40) – (6.43). The nuclear Fock operator in the AO basis is given by

$$F^{p}_{\mu'\nu'} = \frac{1}{S^{\text{RXCHF-ne}}}\left(\sum_{\mu"\nu"} P^{e_1}_{\mu"\nu"}\Gamma_1\left(\mu',\nu';\mu",\nu"\right) + \sum_{\mu\nu}\sum_{\mu"\nu"} P^{e_1}_{\mu"\nu"}P^{\text{tot}}_{\mu\nu}\Gamma_2\left(\mu',\nu';\mu",\nu";\mu,\nu\right)\right) - \frac{E_G}{S^{\text{RXCHF-ne}}}\sum_{\mu"\nu"} P^{e_1}_{\mu"\nu"}\Gamma_{1s}\left(\mu',\nu';\mu",\nu"\right). \tag{E.35}$$

The special electronic Fock operator in the AO basis is given by

$$F^{e_1}_{\mu"\nu"} = \frac{1}{S^{\text{RXCHF-ne}}}\left(\sum_{\mu'\nu'} P^{p}_{\mu'\nu'}\Gamma_1\left(\mu',\nu';\mu",\nu"\right) + \sum_{\mu\nu}\sum_{\mu'\nu'} P^{p}_{\mu'\nu'}P^{\text{tot}}_{\mu\nu}\Gamma_2\left(\mu',\nu';\mu",\nu";\mu,\nu\right)\right) - \frac{E_G}{S^{\text{RXCHF-ne}}}\sum_{\mu'\nu'} P^{p}_{\mu'\nu'}\Gamma_{1s}\left(\mu',\nu';\mu",\nu"\right). \tag{E.36}$$

The regular electronic Fock operator for α-spin orbitals in the AO basis is given by

$$F^{\alpha}_{\mu_1\nu_1} = H^{\text{core}}_{\mu_1\nu_1} + \sum_{\mu_2\nu_2}\left(P^{\text{tot}}_{\mu_2\nu_2}\Gamma^{J}_{ee}\left(\mu_1,\nu_1;\mu_2,\nu_2\right) - P^{\alpha}_{\mu_2\nu_2}\Gamma^{K}_{ee}\left(\mu_1,\nu_1;\mu_2,\nu_2\right)\right) + \frac{1}{S^{\text{RXCHF-ne}}}\sum_{\mu'\nu'}\sum_{\mu"\nu"} P^{p}_{\mu'\nu'}P^{e_1}_{\mu"\nu"}\Gamma_2\left(\mu',\nu';\mu",\nu",\mu_1,\nu_1\right). \tag{E.37}$$

The regular electronic Fock operator for β-spin orbitals in the AO basis is given by

$$F^{\beta}_{\mu_1\nu_1} = H^{\text{core}}_{\mu_1\nu_1} + \sum_{\mu_2\nu_2}\left(P^{\text{tot}}_{\mu_2\nu_2}\Gamma^{J}_{ee}\left(\mu_1,\nu_1;\mu_2,\nu_2\right) - P^{\beta}_{\mu_2\nu_2}\Gamma^{K}_{ee}\left(\mu_1,\nu_1;\mu_2,\nu_2\right)\right) + \frac{1}{S^{\text{RXCHF-ne}}}\sum_{\mu'\nu'}\sum_{\mu"\nu"} P^{p}_{\mu'\nu'}P^{e_1}_{\mu"\nu"}\Gamma_2\left(\mu',\nu';\mu",\nu",\mu_1,\nu_1\right). \tag{E.38}$$

Most integrals required for the evaluation of these quantities are defined in the previous appendix, and the new integrals appearing are defined as

$$\Gamma_2\left(\mu',\nu';\mu",\nu";\mu,\nu\right) = \left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu"}\phi^{e}_{\mu} \middle| \tilde{\Omega}_2 \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu"}\phi^{e}_{\nu} \right\rangle, \tag{E.39}$$

$$\Gamma^{J}_{ee}\left(\mu_1,\nu_1;\mu_2,\nu_2\right)=\left\langle \phi^{e}_{\mu_1}\phi^{e}_{\mu_2}\left|\frac{1}{r_{12}}\right|\phi^{e}_{\nu_1}\phi^{e}_{\nu_2}\right\rangle, \tag{E.40}$$

$$\Gamma^{K}_{ee}\left(\mu_1,\nu_1;\mu_2,\nu_2\right)=\left\langle \phi^{e}_{\mu_1}\phi^{e}_{\mu_2}\left|\frac{1}{r_{12}}\right|\phi^{e}_{\nu_2}\phi^{e}_{\nu_1}\right\rangle. \tag{E.41}$$

E.4. RXCHF-ae Fock Expressions

In this section, we report expressions for the Fock operators in the spin-orbital, spatial-orbital, and atomic-orbital bases for the RXCHF-ae method.

1. Spin-orbital basis

The quantum nuclear Fock operator is obtained by varying the energy in Eq. (6.30) with respect to χ^{p}, leading to

$$\begin{aligned} f^{p}\left(\mathbf{x}^{p}\right)=&\frac{1}{S^{\text{RXCHF-ne}}}\left(\left\langle \chi^{e}_{1}\left|\tilde{\Omega}_1\right|\chi^{e}_{1}\right\rangle+\sum_{a=2}^{N}\left(\left\langle \chi^{e}_{1}\chi^{e}_{a}\left|\tilde{\Omega}_2\right|\chi^{e}_{1}\chi^{e}_{a}\right\rangle-\left\langle \chi^{e}_{1}\chi^{e}_{a}\left|\tilde{\Omega}_{\text{ex}}\right|\chi^{e}_{a}\chi^{e}_{1}\right\rangle\right)\right)\\ &-\frac{E_G+E_{\text{ex}}}{S^{\text{RXCHF-ne}}}\left\langle \chi^{e}_{1}\left|g^{2}\right|\chi^{e}_{1}\right\rangle. \end{aligned} \tag{E.42}$$

The special electronic Fock operator is obtained by varying the energy in Eq. (6.30) with respect to χ^{e}_{1}, leading to

$$\begin{aligned} f^{e_1}\left(\mathbf{x}^{e}_{1}\right)=&\frac{1}{S^{\text{RXCHF-ne}}}\left(\left\langle \chi^{p}\left|\tilde{\Omega}_1\right|\chi^{p}\right\rangle+\sum_{a=2}^{N}\left(\left\langle \chi^{p}\chi^{e}_{a}\left|\tilde{\Omega}_2\right|\chi^{p}\chi^{e}_{a}\right\rangle-\left\langle \chi^{p}\chi^{e}_{a}\left|\tilde{\Omega}_{\text{ex}}P_2\left(1,2\right)\right|\chi^{p}\chi^{e}_{a}\right\rangle\right)\right)\\ &-\frac{E_G+E_{\text{ex}}}{S^{\text{RXCHF-ne}}}\left\langle \chi^{p}\left|g^{2}\right|\chi^{p}\right\rangle. \end{aligned} \tag{E.43}$$

The regular electronic Fock operator is obtained by varying the energy in Eq. (6.30) with respect to some χ^{e}_{μ}, $\mu\in\{2,\ldots N\}$, leading to

$$f^{e}\left(\mathbf{x}_2^{e}\right)=h^{e}\left(2\right)+\sum_{a=2}^{N}\left(J_a\left(2\right)-K_a\left(2\right)\right)+\frac{1}{S^{\text{RXCHF-ne}}}\left(\left\langle\chi^{p}\chi_1^{e}\left|\tilde{\Omega}_2\right|\chi^{p}\chi_1^{e}\right\rangle-\left\langle\chi^{p}\chi_1^{e}\left|\tilde{\Omega}_{\text{ex}}P_2\left(1,2\right)\right|\chi^{p}\chi_1^{e}\right\rangle\right). \tag{E.44}$$

2. Spatial-orbital basis

Next we integrate the RXCHF-ae Fock operators in Eqs. (E.42) - (E.44) over spin to obtain the spatial Fock operators. The quantum nuclear Fock operator is given by

$$f^{p}\left(\mathbf{r}^{p}\right)=\frac{1}{S^{\text{RXCHF-ne}}}\left(\left\langle\psi_1^{e}\left|\tilde{\Omega}_1\right|\psi_1^{e}\right\rangle+\sum_{a}^{N_\alpha}\left[\left\langle\psi_1^{e}\psi_a^{\alpha}\left|\tilde{\Omega}_2\right|\psi_1^{e}\psi_a^{a}\right\rangle-\left\langle\psi_1^{e}\psi_a^{\alpha}\left|\tilde{\Omega}_{\text{ex}}\right|\psi_a^{\alpha}\psi_1^{e}\right\rangle\right]+\sum_{a}^{N_\beta}\left\langle\psi_1^{e}\psi_a^{\beta}\left|\tilde{\Omega}_2\right|\psi_1^{e}\psi_a^{\beta}\right\rangle\right)-\frac{E_G+E_{\text{ex}}}{S^{\text{RXCHF-ne}}}\left\langle\chi_1^{e}\left|g^2\right|\chi_1^{e}\right\rangle. \tag{E.45}$$

The special electronic Fock operator is given by

$$f^{e_1}\left(\mathbf{r}_1^{e}\right)=\frac{1}{S^{\text{RXCHF-ne}}}\left(\left\langle\psi^{p}\left|\tilde{\Omega}_1\right|\psi^{p}\right\rangle+\sum_{a}^{N_\alpha}\left[\left\langle\psi^{p}\psi_a^{\alpha}\left|\tilde{\Omega}_2\right|\psi^{p}\psi_a^{\alpha}\right\rangle-\left\langle\psi^{p}\psi_a^{\alpha}\left|\tilde{\Omega}_{\text{ex}}P_2\left(1,2\right)\right|\psi^{p}\psi_a^{\alpha}\right\rangle\right]+\sum_{a}^{N_\beta}\left\langle\psi^{p}\psi_a^{\beta}\left|\tilde{\Omega}_2\right|\psi^{p}\psi_a^{\beta}\right\rangle\right)-\frac{E_G+E_{\text{ex}}}{S^{\text{RXCHF-ne}}}\left\langle\chi^{p}\left|g^2\right|\chi^{p}\right\rangle. \tag{E.46}$$

The regular electronic Fock operator for α-spin orbitals is given by

$$f^{\alpha}\left(\mathbf{x}_2^{e}\right)=h^{e}\left(2\right)+\sum_{a}^{N_\alpha}\left(J_a^{\alpha}\left(2\right)-K_a^{\alpha}\left(2\right)\right)+\sum_{a}^{N_\beta}J_a^{\beta}\left(2\right)+\frac{1}{S^{\text{RXCHF-ne}}}\left(\left\langle\psi^{p}\psi_1^{e}\left|\tilde{\Omega}_2\right|\psi^{p}\psi_1^{e}\right\rangle-\left\langle\psi^{p}\psi_1^{e}\left|\tilde{\Omega}_{\text{ex}}P_2\left(1,2\right)\right|\psi^{p}\psi_1^{e}\right\rangle\right). \tag{E.47}$$

The regular electronic Fock operator for β-spin orbitals is given by

$$f^{\beta}\left(\mathbf{x}_2^{e}\right)=h^{e}\left(2\right)+\sum_{a}^{N_\beta}\left(J_a^{\beta}\left(2\right)-K_a^{\beta}\left(2\right)\right)+\sum_{a}^{N_\alpha}J_a^{\alpha}\left(2\right)+\frac{1}{S^{\text{RXCHF-ne}}}\left\langle\psi^{p}\psi_1^{e}\left|\tilde{\Omega}_2\right|\psi^{p}\psi_1^{e}\right\rangle. \tag{E.48}$$

3. Atomic-orbital basis

We now expand the spatial orbitals in the AO bases, as in Eqs. (6.40) – (6.43). The nuclear Fock operator in the AO basis is given by

$$\begin{aligned} F^{p}_{\mu'\nu'} = \frac{1}{S^{\text{RXCHF-ne}}}\Bigg(& \sum_{\mu"\nu"} P^{e_1}_{\mu"\nu"}\Gamma_1\left(\mu',\nu';\mu",\nu"\right) + \sum_{\mu\nu}\sum_{\mu"\nu"} P^{e_1}_{\mu"\nu"} \\ & \times\left[P^{\text{tot}}_{\mu\nu}\Gamma_2\left(\mu',\nu';\mu",\nu";\mu,\nu\right) - P^{\alpha}_{\mu\nu}\Gamma_{\text{ex}}\left(\mu',\nu';\mu",\nu";\mu,\nu\right)\right]\Bigg) \\ & - \frac{E_G + E_{\text{ex}}}{S^{\text{RXCHF-ne}}}\sum_{\mu"\nu"} P^{e_1}_{\mu"\nu"}\Gamma_{1s}\left(\mu',\nu';\mu",\nu"\right). \end{aligned} \tag{E.49}$$

The special electronic Fock operator in the AO basis is given by

$$\begin{aligned} F^{e_1}_{\mu"\nu"} = \frac{1}{S^{\text{RXCHF-ne}}}\Bigg(& \sum_{\mu'\nu'} P^{p}_{\mu'\nu'}\Gamma_1\left(\mu',\nu';\mu",\nu"\right) + \sum_{\mu\nu}\sum_{\mu'\nu'} P^{p}_{\mu'\nu'} \\ & \times\left[P^{\text{tot}}_{\mu\nu}\Gamma_2\left(\mu',\nu';\mu",\nu";\mu,\nu\right) - P^{\alpha}_{\mu\nu}\Gamma_{\text{ex}}\left(\mu',\nu';\mu",\nu";\mu,\nu\right)\right]\Bigg) \\ & - \frac{E_G + E_{\text{ex}}}{S^{\text{RXCHF-ne}}}\sum_{\mu'\nu'} P^{p}_{\mu'\nu'}\Gamma_{1s}\left(\mu',\nu';\mu",\nu"\right). \end{aligned} \tag{E.50}$$

The regular electronic Fock operator for α-spin orbitals in the AO basis is given by

$$\begin{aligned} F^{\alpha}_{\mu_1\nu_1} = & H^{\text{core}}_{\mu_1\nu_1} + \sum_{\mu_2\nu_2}\left(P^{\text{tot}}_{\mu_2\nu_2}\Gamma^{J}_{ee}\left(\mu_1,\nu_1;\mu_2,\nu_2\right) - P^{\alpha}_{\mu_2\nu_2}\Gamma^{K}_{ee}\left(\mu_1,\nu_1;\mu_2,\nu_2\right)\right) \\ & + \frac{1}{S^{\text{RXCHF-ne}}}\sum_{\mu'\nu'}\sum_{\mu"\nu"} P^{p}_{\mu'\nu'}P^{e_1}_{\mu"\nu"}\left(\Gamma_2\left(\mu',\nu';\mu",\nu",\mu_1,\nu_1\right) - \Gamma_{\text{ex}}\left(\mu',\nu';\mu",\nu",\mu_1,\nu_1\right)\right). \end{aligned} \tag{E.51}$$

The regular electronic Fock operator for β-spin orbitals in the AO basis is given by

$$\begin{aligned} F^{\beta}_{\mu_1\nu_1} = & H^{\text{core}}_{\mu_1\nu_1} + \sum_{\mu_2\nu_2}\left(P^{\text{tot}}_{\mu_2\nu_2}\Gamma^{J}_{ee}\left(\mu_1,\nu_1;\mu_2,\nu_2\right) - P^{\beta}_{\mu_2\nu_2}\Gamma^{K}_{ee}\left(\mu_1,\nu_1;\mu_2,\nu_2\right)\right) \\ & + \frac{1}{S^{\text{RXCHF-ne}}}\sum_{\mu'\nu'}\sum_{\mu"\nu"} P^{p}_{\mu'\nu'}P^{e_1}_{\mu"\nu"}\left(\Gamma_2\left(\mu',\nu';\mu",\nu",\mu_1,\nu_1\right) - \Gamma_{\text{ex}}\left(\mu',\nu';\mu",\nu",\mu_1,\nu_1\right)\right). \end{aligned} \tag{E.52}$$

Most integrals required for the evaluation of these quantities are defined in the previous appendix, and the new integral appearing is defined as

$$\Gamma_{\text{ex}}\left(\mu',\nu';\mu",\nu";\mu,\nu\right) = \left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu"}\phi^{e}_{\mu} \middle| \tilde{\Omega}_{\text{ex}} \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu}\phi^{e}_{\nu"} \right\rangle. \tag{E.53}$$

E.5. One-Electron RXCHF-fe Energy

In this section, we present the detailed derivation of the one-electron RXCHF-fe energy, including the operator formalism used for the computation of this and other energy contributions. We define operators by their action on the $N \times N$ Slater determinant (SD) that appears in the RXCHF-fe wavefunction:

$$\Psi^e(1,\ldots,N,p) = \frac{1}{\sqrt{N}} \begin{vmatrix} \chi_1^e(1)\, g(1,p) & \chi_2^e(1) & \cdots & \chi_N^e(1) \\ \chi_1^e(2)\, g(2,p) & \chi_2^e(2) & \cdots & \chi_N^e(2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1^e(N)\, g(N,p) & \chi_2^e(N) & \ldots & \chi_N^e(N) \end{vmatrix}, \tag{E.54}$$

where the complete RXCHF-fe wavefunction is $\Psi^{\text{RXCHF-fe}}(1,\ldots,N,p) = \chi^p(p)\Psi^e(1,\ldots,N,p)$.

We define a set of *column operators*, $\{b_i : 1 \le i \le N\}$, by their action on Ψ^e:

$$b_i\Psi^e \equiv \left(b_i\Psi^e\right)(1,\ldots,i-1,i+1,\ldots N) = \frac{(-1)^{i+1}}{\sqrt{(N-1)!}} \begin{vmatrix} \chi_2^e(1) & \cdots & \chi_N^e(1) \\ \vdots & & \vdots \\ \chi_2^e(i-1) & \cdots & \chi_N^e(i-1) \\ \chi_2^e(i+1) & \cdots & \chi_N^e(i+1) \\ \vdots & & \vdots \\ \chi_2^e(N) & \cdots & \chi_N^e(N) \end{vmatrix}, \tag{E.55}$$

i.e., $b_i\Psi^e$ is obtained from Ψ^e by deleting the first column and the i^{th} row, removing the i^{th}-coordinate dependence as well as the dependence on the proton coordinate. Clearly,

$$\Psi^e = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \chi_1^e(i)\, g(i,p)\left(b_i\Psi^e\right) \tag{E.56}$$

is the expansion of the $N \times N$ SD Ψ^e along the first column.

We also define a set of *row operators*, $\left\{a_k^{(j)} : 2 \le k \le N, 1 \le j \le N\right\}$, by their action on $b_i\Psi^e$ for any $i \in \{1,\ldots,N\} \setminus \{j\}$:

$$a_k^{(j)} b_i \Psi^e \equiv \left(a_k^{(j)} b_i \Psi^e\right)(1,\ldots,i-1,i+1,\ldots,j-1,j+1,\ldots N)$$
$$= \frac{\theta(i,j,k)}{\sqrt{(N-2)!}} \begin{vmatrix} \chi_2^e(1) & \cdots & \chi_{k-1}^e(1) & \chi_{k+1}^e(1) & \cdots & \chi_N^e(1) \\ \vdots & & \vdots & \vdots & & \vdots \\ \chi_2^e(i-1) & \cdots & \chi_{k-1}^e(i-1) & \chi_{k+1}^e(i-1) & \cdots & \chi_N^e(i-1) \\ \chi_2^e(i+1) & \cdots & \chi_{k-1}^e(i+1) & \chi_{k+1}^e(i+1) & \cdots & \chi_N^e(i+1) \\ \vdots & & \vdots & \vdots & & \vdots \\ \chi_2^e(j-1) & \cdots & \chi_{k-1}^e(j-1) & \chi_{k+1}^e(j-1) & \cdots & \chi_N^e(j-1) \\ \chi_2^e(j+1) & \cdots & \chi_{k-1}^e(j+1) & \chi_{k+1}^e(j+1) & \cdots & \chi_N^e(j+1) \\ \vdots & & \vdots & \vdots & & \vdots \\ \chi_2^e(N) & \cdots & \chi_{k-1}^e(N) & \chi_{k+1}^e(N) & \cdots & \chi_N^e(N) \end{vmatrix}, \quad \text{(E.57)}$$

where we have written the SD as if $j > i$ for convenience (reverse i and j in the above SD if $j < i$). In this expression, $\theta(i,j,k)$ is a phase factor defined by

$$\theta(i,j,k) = (-1)^{i+1} \begin{cases} (-1)^{j+k} & j > i \\ (-1)^{j+k+1} & j < i \end{cases}. \quad \text{(E.58)}$$

With this definition, $a_k^{(j)} b_i \Psi^e$ is obtained from $b_i\Psi^e$ by deleting the k^{th} column and the j^{th} row, removing the j^{th}-coordinate dependence. Similar to above, we have for any $j \in \{1,\ldots,N\} \setminus \{i\}$,

$$b_i \Psi^e = \frac{1}{\sqrt{N-1}} \sum_{k=2}^{N} \chi_k^e(j) \left(a_k^{(j)} b_i \Psi^e\right) \quad \text{(E.59)}$$

as the expansion of the $(N-1)\times(N-1)$ SD $b_i\Psi^e$ along the j^{th} row. The row and column operators presented here are similar to those introduced by Hinze in his multi-configuration self-consistent field (MCSCF) formulation.[2]

With these definitions, we can evaluate the expectation value over general one-electron operators. We compute

$$\begin{aligned}\left\langle \Psi \middle| \hat{O}_1^e \middle| \Psi \right\rangle &= \left\langle \chi^p(p)\Psi^e(1,\ldots,N,p) \middle| \sum_{i=1}^{N} h_e(i) \middle| \chi^p(p)\Psi^e(1,\ldots,N,p) \right\rangle \\ &= N\left\langle \chi^p(p)\Psi^e(1,\ldots,N,p) \middle| h_e(1) \middle| \chi^p(p)\Psi^e(1,\ldots,N,p) \right\rangle \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N}\left\langle \chi^p(p)\chi_1^e(i)g(i,p)\left(b_i\Psi^e\right) \middle| h_e(1) \middle| \chi^p(p)\chi_1^e(j)g(j,p)\left(b_j\Psi^e\right) \right\rangle.\end{aligned} \tag{E.60}$$

The $i=j$ terms are given by

$$\begin{aligned}&\left\langle \chi^p(p)\chi_1^e(1)g(1,p)\left(b_1\Psi^e\right) \middle| h_e(1) \middle| \chi^p(p)\chi_1^e(1)g(1,p)\left(b_1\Psi^e\right) \right\rangle \\ &+\sum_{i=2}^{N}\left\langle \chi^p(p)\chi_1^e(i)g(i,p)\left(b_i\Psi^e\right) \middle| h_e(1) \middle| \chi^p(p)\chi_1^e(i)g(i,p)\left(b_i\Psi^e\right) \right\rangle \\ &=\left\langle \chi^p(p)\chi_1^e(1) \middle| g(1,p)h_e(1)g(1,p) \middle| \chi^p(p)\chi_1^e(1) \right\rangle \left\langle \Psi^e \middle| b_1^\dagger b_1 \middle| \Psi^e \right\rangle \\ &+\sum_{i=2}^{N}\left\langle \chi^p(p)\chi_1^e(i) \middle| g(i,p)^2 \middle| \chi^p(p)\chi_1^e(i) \right\rangle \left\langle \Psi^e \middle| b_i^\dagger h^e(1) b_i \middle| \Psi^e \right\rangle.\end{aligned} \tag{E.61}$$

The matrix elements over column operators can be computed readily using the Slater rules for $(N-1)\times(N-1)$ SDs. The resulting expression for the $i=j$ terms is

$$\begin{aligned}&\left\langle \chi^p(p)\chi_1^e(1) \middle| g(1,p)h_e(1)g(1,p) \middle| \chi^p(p)\chi_1^e(1) \right\rangle \\ &+\left\langle \chi^p(p)\chi_1^e(1) \middle| g(1,p)^2 \middle| \chi^p(p)\chi_1^e(1) \right\rangle \sum_{a=2}^{N}\left\langle \chi_a^e(1) \middle| h^e(1) \middle| \chi_a^e(1) \right\rangle.\end{aligned} \tag{E.62}$$

The $i\neq j$ terms are given by

$$
\begin{aligned}
&\sum_{j=2}^{N}\left\langle \chi^{p}(p)\,\chi_{1}^{e}(1)\,g(1,p)\left(b_{1}\Psi^{e}\right)\middle|h_{e}(1)\middle|\chi^{p}(p)\,\chi_{1}^{e}(j)\,g(j,p)\left(b_{j}\Psi^{e}\right)\right\rangle \\
&+\sum_{i=2}^{N}\left\langle \chi^{p}(p)\,\chi_{1}^{e}(i)\,g(i,p)\left(b_{i}\Psi^{e}\right)\middle|h_{e}(1)\middle|\chi^{p}(p)\,\chi_{1}^{e}(1)\,g(1,p)\left(b_{1}\Psi^{e}\right)\right\rangle \\
&+\sum_{i=2}^{N}\sum_{j\neq i}^{N}\left\langle \chi^{p}(p)\,\chi_{1}^{e}(i)\,g(i,p)\left(b_{i}\Psi^{e}\right)\middle|h_{e}(1)\middle|\chi^{p}(p)\,\chi_{1}^{e}(j)\,g(j,p)\left(b_{j}\Psi^{e}\right)\right\rangle \\
&=\frac{1}{N-1}\Bigg(\sum_{j=2}^{N}\sum_{k=2}^{N}\sum_{l=2}^{N} \\
&\left\langle \chi^{p}(p)\,\chi_{1}^{e}(1)\,g(1,p)\,\chi_{k}^{e}(j)\left(a_{k}^{(j)}b_{1}\Psi^{e}\right)\middle|h_{e}(1)\middle|\chi^{p}(p)\,\chi_{1}^{e}(j)\,g(j,p)\,\chi_{l}^{e}(1)\left(a_{l}^{(1)}b_{j}\Psi^{e}\right)\right\rangle \\
&+\sum_{i=2}^{N}\sum_{k=2}^{N}\sum_{l=2}^{N} \\
&\left\langle \chi^{p}(p)\,\chi_{1}^{e}(i)\,g(i,p)\,\chi_{k}^{e}(1)\left(a_{k}^{(1)}b_{i}\Psi^{e}\right)\middle|h_{e}(1)\middle|\chi^{p}(p)\,\chi_{1}^{e}(1)\,g(1,p)\,\chi_{l}^{e}(i)\left(a_{l}^{(i)}b_{1}\Psi^{e}\right)\right\rangle \\
&+\sum_{i=2}^{N}\sum_{j\neq i}^{N}\sum_{k=2}^{N}\sum_{l=2}^{N} \\
&\left\langle \chi^{p}(p)\,\chi_{1}^{e}(i)\,g(i,p)\,\chi_{k}^{e}(j)\left(a_{k}^{(j)}b_{i}\Psi^{e}\right)\middle|h_{e}(1)\middle|\chi^{p}(p)\,\chi_{1}^{e}(j)\,g(j,p)\,\chi_{l}^{e}(i)\left(a_{l}^{(i)}b_{j}\Psi^{e}\right)\right\rangle\Bigg) \\
&=\frac{1}{N-1}\Bigg(\sum_{j=2}^{N}\sum_{k=2}^{N}\sum_{l=2}^{N} \\
&\left\langle \chi^{p}(p)\,\chi_{1}^{e}(1)\,\chi_{k}^{e}(j)\middle|g(1,p)\,h_{e}(1)\,g(j,p)\middle|\chi^{p}(p)\,\chi_{1}^{e}(j)\,\chi_{l}^{e}(1)\right\rangle\left\langle \Psi^{e}\middle|b_{1}^{\dagger}a_{k}^{(j)\dagger}a_{l}^{(1)}b_{j}\middle|\Psi^{e}\right\rangle \\
&+\sum_{i=2}^{N}\sum_{k=2}^{N}\sum_{l=2}^{N} \\
&\left\langle \chi^{p}(p)\,\chi_{1}^{e}(i)\,\chi_{k}^{e}(1)\middle|g(i,p)\,h_{e}(1)\,g(1,p)\middle|\chi^{p}(p)\,\chi_{1}^{e}(1)\,\chi_{l}^{e}(i)\right\rangle\left\langle \Psi^{e}\middle|b_{i}^{\dagger}a_{k}^{(1)\dagger}a_{l}^{(i)}b_{1}\middle|\Psi^{e}\right\rangle \\
&+\sum_{i=2}^{N}\sum_{j\neq i}^{N}\sum_{k=2}^{N}\sum_{l=2}^{N} \\
&\left\langle \chi^{p}(p)\,\chi_{1}^{e}(i)\,\chi_{k}^{e}(j)\middle|g(i,p)\,g(j,p)\middle|\chi^{p}(p)\,\chi_{1}^{e}(j)\,\chi_{l}^{e}(i)\right\rangle\left\langle \Psi^{e}\middle|b_{i}^{\dagger}a_{k}^{(j)\dagger}h^{e}(1)\,a_{l}^{(i)}b_{j}\middle|\Psi^{e}\right\rangle\Bigg).
\end{aligned}
\tag{E.63}
$$

Again, the matrix elements over column operators can be computed readily using the Slater rules for $(N-2)\times(N-2)$ SDs (the last term is first split into $k=l$ and $k\neq l$ contributions). The resulting expression for the $i\neq j$ terms is

$$
\begin{aligned}
&-\sum_{a=2}^{N}\left\langle \chi^{p}(p)\,\chi_{1}^{e}(1)\,\chi_{a}^{e}(2)\middle|g(1,p)\,h_{e}(1)\,g(2,p)\middle|\chi^{p}(p)\,\chi_{1}^{e}(2)\,\chi_{a}^{e}(1)\right\rangle \\
&-\sum_{a=2}^{N}\left\langle \chi^{p}(p)\,\chi_{1}^{e}(2)\,\chi_{a}^{e}(1)\middle|g(2,p)\,h_{e}(1)\,g(1,p)\middle|\chi^{p}(p)\,\chi_{1}^{e}(1)\,\chi_{a}^{e}(2)\right\rangle \\
&-\sum_{a=2}^{N}\sum_{b\neq a}^{N}\left\langle \chi^{p}(p)\,\chi_{1}^{e}(1)\,\chi_{a}^{e}(2)\middle|g(1,p)\,g(2,p)\middle|\chi^{p}(p)\,\chi_{1}^{e}(2)\,\chi_{a}^{e}(1)\right\rangle\left\langle \chi_{b}^{e}(1)\middle|h^{e}(1)\middle|\chi_{b}^{e}(1)\right\rangle \\
&+\sum_{a=2}^{N}\sum_{b\neq a}^{N}\left\langle \chi^{p}(p)\,\chi_{1}^{e}(1)\,\chi_{a}^{e}(2)\middle|g(1,p)\,g(2,p)\middle|\chi^{p}(p)\,\chi_{1}^{e}(2)\,\chi_{b}^{e}(1)\right\rangle\left\langle \chi_{b}^{e}(1)\middle|h^{e}(1)\middle|\chi_{a}^{e}(1)\right\rangle .
\end{aligned}
\tag{E.64}
$$

Combining Eqs. (E.62) and (E.64) and removing the redundant restriction on the double summation, the expectation value over the core electron operators is evaluated to be

$$
\begin{aligned}
\left\langle \Psi\middle|\hat{O}_{1}^{e}\middle|\Psi\right\rangle &= \left\langle \chi^{p}(p)\,\chi_{1}^{e}(1)\middle|g(1,p)\,h_{e}(1)\,g(1,p)\middle|\chi^{p}(p)\,\chi_{1}^{e}(1)\right\rangle \\
&\quad+\sum_{a=2}^{N}\Bigg[\left\langle \chi^{p}(p)\,\chi_{1}^{e}(1)\middle|g(1,p)^{2}\middle|\chi^{p}(p)\,\chi_{1}^{e}(1)\right\rangle\left\langle \chi_{a}^{e}(1)\middle|h^{e}(1)\middle|\chi_{a}^{e}(1)\right\rangle \\
&\quad-\left\langle \chi^{p}(p)\,\chi_{1}^{e}(1)\,\chi_{a}^{e}(2)\middle|g(1,p)\,h_{e}(1)\,g(2,p)\middle|\chi^{p}(p)\,\chi_{a}^{e}(1)\,\chi_{1}^{e}(2)\right\rangle \\
&\quad-\left\langle \chi^{p}(p)\,\chi_{a}^{e}(1)\,\chi_{1}^{e}(2)\middle|g(2,p)\,h_{e}(1)\,g(1,p)\middle|\chi^{p}(p)\,\chi_{1}^{e}(1)\,\chi_{a}^{e}(2)\right\rangle\Bigg] \\
&\quad+\sum_{a=2}^{N}\sum_{b=2}^{N} \\
&\Bigg[\left\langle \chi^{p}(p)\,\chi_{1}^{e}(1)\,\chi_{a}^{e}(2)\middle|g(1,p)\,g(2,p)\middle|\chi^{p}(p)\,\chi_{b}^{e}(1)\,\chi_{1}^{e}(2)\right\rangle\left\langle \chi_{b}^{e}(1)\middle|h^{e}(1)\middle|\chi_{a}^{e}(1)\right\rangle \\
&\quad-\left\langle \chi^{p}(p)\,\chi_{1}^{e}(1)\,\chi_{a}^{e}(2)\middle|g(1,p)\,g(2,p)\middle|\chi^{p}(p)\,\chi_{a}^{e}(1)\,\chi_{1}^{e}(2)\right\rangle\left\langle \chi_{b}^{e}(1)\middle|h^{e}(1)\middle|\chi_{b}^{e}(1)\right\rangle\Bigg].
\end{aligned}
\tag{E.65}
$$

E.6. Orthogonalization Scheme

In this section, we discuss an alternative orthogonalization scheme that was used in some of the calculations. For the implementation used in Chapter 7 on positronic systems, these two orthogonalization schemes are mathematically equivalent. In this alternative scheme, we also use a modification of the orthogonality-constrained basis set expansion (OCBSE) method[3] in order to ensure that the geminal-coupled electronic orbital remains orthogonal to the regular electronic orbitals. The only difference between the present scheme and the scheme presented in

the main paper is the choice of the basis in which to expand the geminal-coupled electronic orbital, namely $B^{(i)}$. We adopt the same notation given in the main paper and highlight the differences between the two approaches below.

In this scheme, the basis to which the modified Fock equation for the geminal-coupled electronic orbital is transformed is defined as

$$B^{(i)} = \left\{\tilde{\psi}^{(i-1)}\right\} \cup \left(\left\{ \psi_a^{(i)} : N_\alpha < a \le N^{\text{ebf}} \right\} \setminus \left\{ \psi_b^{(i)} \right\} \right) \text{ for some } N_\alpha < b \le N^{\text{ebf}}. \quad \text{(E.66)}$$

Here $\tilde{\psi}^{(i-1)}$ is the normalized vector along

$$\psi^{(i-1)} - \sum_{a \in I} \left\langle \psi^{(i-1)} \middle| \psi_a^{(i)} \right\rangle \psi_a^{(i)}, \quad I = \left\{1, \ldots N^{\text{ebf}}\right\} \setminus \{b\}, \quad \text{(E.67)}$$

which is the projection of $\psi^{(i-1)}$ obtained via Gram-Schmidt orthogonalization with respect to the regular alpha electronic orbitals. The orbital removed from $B^{(i)}$ in Eq. (E.66) is chosen as the virtual orbital that is in maximal coincidence with the occupied geminal-coupled electronic orbital $\psi^{(i-1)}$ of the previous iteration. Specifically, $\psi_b^{(i)}$ is chosen such that

$$\left\langle \psi^{(i-1)} \middle| \psi_b^{(i)} \right\rangle > \left\langle \psi^{(i-1)} \middle| \psi_a^{(i)} \right\rangle \quad \forall a \in \left\{ N_\alpha + 1, \ldots, N^{\text{ebf}} \right\} \setminus \{b\}. \quad \text{(E.68)}$$

The justification for this definition is as follows. We remove from the orthogonal set of the virtual regular alpha electronic orbitals of the current iteration the virtual orbital $\psi_b^{(i)}$ that overlaps the most with $\psi^{(i-1)}$. We then add to $B^{(i)}$ the projected function $\tilde{\psi}^{(i-1)}$, which by construction is orthogonal to the rest of the occupied and virtual regular alpha electronic orbitals. Therefore, $B^{(i)}$ is an orthogonal set containing a projected form of the geminal-coupled electronic orbital from the previous iteration, and the space spanned by its elements is orthogonal to the space spanned by the occupied regular alpha electronic orbitals of the current iteration.

This scheme was adopted initially to more closely parallel the original OCBSE procedure implemented by Goddard *et al.*, where orthogonality was enforced between both subsets of electronic orbitals.[3] In our modified scheme presented in the main paper, orthogonality is only enforced in the Fock procedure for the geminal-coupled electronic orbital, and the optimization with respect to the regular electronic orbitals is carried through at each step of the SCF procedure without constraint.

E.7. Operator Symmetrizations

In this section, we briefly discuss the operator symmetrization procedure used for the RXCHF methods. In order to generate the compact Fock equations presented in the main paper, we apply operator symmetrizations similar to those used in the XCHF methods.[1,4] The replacement of the operators given in the main paper with their symmetric forms given below leads to identical energy expressions because the summations over terms involving these operators is invariant to the symmetrizations used. The procedure for RXCHF is not identical, however, since the geminal-coupled electronic orbital is not equivalent to the regular electronic orbitals. Therefore, we must be specific about which indices of each operator we would like to symmetrize, being careful not to symmetrize all terms involving the geminal-coupled electronic orbital. In what follows, we define the specific symmetric forms of the operators used in the RXCHF methods.

In the RXCHF-fe method, the operator symmetrizations are as follows:

$$\tilde{\Omega}_1(p,1) = \Omega_1(p,1), \tag{E.69}$$

$$\tilde{\Omega}_2^{(1)}(p,1,2) = \Omega_2^{(1)}(p,1,2), \tag{E.70}$$

$$\tilde{\Omega}_2^{(2)}\left(p,1,2\right)=\frac{1}{2}\left(\Omega_2^{(2)}\left(p,1,2\right)+\Omega_2^{(2)}\left(p,2,1\right)\right), \tag{E.71}$$

$$\tilde{\Omega}_3^{(1)}\left(p,1,2,3\right)=\frac{1}{2}\left(\Omega_3^{(1)}\left(p,1,2,3\right)+\Omega_3^{(1)}\left(p,1,3,2\right)\right), \tag{E.72}$$

$$\tilde{\Omega}_3^{(2)}\left(p,1,2,3\right)=\Omega_3^{(2)}\left(p,1,2,3\right). \tag{E.73}$$

In particular, the inability to symmetrize $\Omega_3^{(2)}\left(p,1,2,3\right)$ leads to the difficulty in expressing the regular electronic Fock operator in a compact form (see Section E.2), but compact expressions for the other Fock operators can still be expressed because terms involving $\Omega_3^{(2)}$ are all summed over the regular electronic indices.

In the RXCHF-ne and RXCHF-ae methods, the additional operator symmetrizations are as follows:

$$\tilde{\Omega}_2\left(p,1,2\right)=\Omega_2\left(p,1,2\right), \tag{E.74}$$

$$\tilde{\Omega}_{\mathrm{ex}}\left(p,1,2\right)=\frac{1}{2}\left(\Omega_{\mathrm{ex}}\left(p,1,2\right)+\Omega_{\mathrm{ex}}\left(p,2,1\right)\right). \tag{E.75}$$

Note that no symmetrization is needed for RXCHF-ne because the geminal contributions are always associated with the geminal-coupled electronic orbital (i.e., there is no exchange of electronic indices).

References

(1) A. Chakraborty, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Physics* **129**, 014101 (2008).
(2) J. Hinze, *Journal of Chemical Physics* **59**, 6424 (1973).
(3) W. J. Hunt, T. H. Dunning, and W. A. Goddard, *Chemical Physics Letters* **3**, 606 (1969).
(4) C. Ko, M. V. Pak, C. Swalina, and S. Hammes-Schiffer, *Journal of Chemical Physics* **135**, 054106 (2011).

Appendix F

Supporting Information for Chapter 7[‡]

F.1. Summary of Contents

In this Appendix, we include supporting information corresponding to the material presented in Chapter 7. We include annihilation rate expressions applicable for the e^+Li system for RXCHF-fe, RXCHF-ae and RXCHF-ne. We also include a discussion of results for the PsH system. In addition, we include the following figures: densities for PsH, for e^+Li, LiPs and e^+LiH using GTG scheme (i), for e^+Li with a larger basis set, and for e^+LiH using GTG scheme (ii) with RXCHF-ae. We also include the following tables: expectation values and annihilation rates for PsH, for e^+Li, LiPs and e^+LiH using GTG scheme (i), and for e^+Li with a larger basis set.

F.2. Annihilation Rate Expressions

In this section, we report expressions for the annihilation rate in the spatial orbital basis using the RXCHF methods for the e^+Li system. These expressions are obtained by evaluating Eq. (7.11) with appropriate action of the spin-projection operator, $\hat{O}^{s}_{e_i,p}$.

[‡]Reproduced with permission from:
A. Sirjoosingh, M. V. Pak, C. Swalina, and S. Hammes-Schiffer, "Reduced explicitly correlated Hartree-Fock approach within the nuclear-electronic orbital framework: Applications to positronic molecular systems," *J. Chem. Phys.* **139**, 034103 (2013). © 2013 American Institute of Physics

As discussed in Section 7.3A, there are two scenarios for positronium annihilation in e^+Li: (i) singlet annihilation arising from the unpaired electron being the opposite spin as the positron, and (ii) triplet annihilation arising from the unpaired electron being the same spin as the positron. The rates in each case are defined in Eqs. (7.13) and (7.14). Without loss of generality, we assume that the spin of the geminal-coupled electronic orbital is α:

$$\chi_1^e\left(\mathbf{r}_1^e,\omega_1^e\right)=\psi_1^e\left(\mathbf{r}_1^e\right)\alpha\left(\omega_1^e\right). \tag{F.1}$$

We also assume that the regular electronic orbitals are partitioned as

$$\left\{\chi_a^e\left(\mathbf{r}_a^e,\omega_a^e\right)\right\}=\left\{\psi_a^\alpha\left(\mathbf{r}_a^e\right)\alpha\left(\omega_a^e\right)\right\}\cup\left\{\psi_a^\beta\left(\mathbf{r}_a^e\right)\beta\left(\omega_a^e\right)\right\}\equiv \mathrm{A}\cup\mathrm{B}, \tag{F.2}$$

where $|\mathrm{A}|=N_\alpha$ and $|\mathrm{B}|=N_\beta$ such that $N_\alpha+N_\beta=N-1$. Note that the index *a* spans a different range in the context of the spatial orbitals: for ψ_a^α, $1\le a\le N_\alpha$, and for ψ_a^β, $1\le a\le N_\beta$.

In the singlet case, the positron has spin β, and the action of the spin-projection operator is given by

$$\left\langle\alpha\left(\omega_1^e\right)\middle|\hat{O}_{e_1,p}^s\middle|\alpha\left(\omega_1^e\right)\right\rangle_{\omega_1^e}=1 \quad\text{and}\quad \left\langle\beta\left(\omega_1^e\right)\middle|\hat{O}_{e_1,p}^s\middle|\beta\left(\omega_1^e\right)\right\rangle_{\omega_1^e}=0. \tag{F.3}$$

In the triplet case, the positron has spin α, and the action of the spin-projection operator is given by

$$\left\langle\beta\left(\omega_1^e\right)\middle|\hat{O}_{e_1,p}^s\middle|\beta\left(\omega_1^e\right)\right\rangle_{\omega_1^e}=1 \quad\text{and}\quad \left\langle\alpha\left(\omega_1^e\right)\middle|\hat{O}_{e_1,p}^s\middle|\alpha\left(\omega_1^e\right)\right\rangle_{\omega_1^e}=0. \tag{F.4}$$

The subscripts following the angular brackets indicate the coordinate(s) of integration.

Herein we report expressions for $\delta_{ep}^{\mathrm{sing}}(\mathbf{r})$ and $\delta_{ep}^{\mathrm{trip}}(\mathbf{r})$, which in practice are numerically integrated to calculate $\langle\delta_{ep}\rangle_{\mathrm{sing}}$ and $\langle\delta_{ep}\rangle_{\mathrm{trip}}$, respectively. Analytical expressions have also been derived for the latter quantities but are not presented here. These analytical expressions were

used for calculations on the non-spherically symmetric e^+LiH system to avoid complications with three-dimensional numerical integration.

1. RXCHF-ne and RXCHF-ae

Since the RXCHF-ae energy is not associated with a particular wavefunction, we evaluate the relevant expectation values using the RXCHF-ne wavefunction. Thus, the corresponding annihilation rate expressions are the same and only differ in the orbitals, which are variationally optimized with different energy expressions.

The singlet contribution is determined by evaluating the integral on the right-hand side of Eq. (7.11) using Eq. (F.3). The resulting expression is $\left\langle \delta_{ep} \right\rangle_{\text{sing}} = \int \mathrm{d}\mathbf{r} \delta_{ep}^{\text{sing}}(\mathbf{r})$, where

$$\delta_{ep}^{\text{sing}}(\mathbf{r}) = \frac{\left|\psi^p(\mathbf{r})\right|^2}{\left\langle \Psi^{\text{RXCHF-ne}} \middle| \Psi^{\text{RXCHF-ne}} \right\rangle} \left(\left|\psi_1^e(\mathbf{r})\right|^2 \left[g(\mathbf{r},\mathbf{r})\right]^2 + \sum_a^{N_\alpha} \left|\psi_a^\alpha(\mathbf{r})\right|^2 \left\langle \psi_1^e \middle| \left[g(\mathbf{r}_1^e,\mathbf{r})\right]^2 \middle| \psi_1^e \right\rangle_{\mathbf{r}_1^e} \right). \quad \text{(F.5)}$$

The triplet contribution is determined by evaluating the integral on the right-hand side of Eq. (7.11) using Eq. (F.4). The resulting expression is $\left\langle \delta_{ep} \right\rangle_{\text{trip}} = \int \mathrm{d}\mathbf{r} \delta_{ep}^{\text{trip}}(\mathbf{r})$, where

$$\delta_{ep}^{\text{trip}}(\mathbf{r}) = \frac{\left|\psi^p(\mathbf{r})\right|^2}{\left\langle \Psi^{\text{RXCHF-ne}} \middle| \Psi^{\text{RXCHF-ne}} \right\rangle} \sum_a^{N_\beta} \left|\psi_a^\beta(\mathbf{r})\right|^2 \left\langle \psi_1^e \middle| \left[g(\mathbf{r}_1^e,\mathbf{r})\right]^2 \middle| \psi_1^e \right\rangle_{\mathbf{r}_1^e}. \quad \text{(F.6)}$$

2. RXCHF-fe

The singlet contribution is determined by evaluating the integral on the right-hand side of Eq. (7.11) using Eq. (F.3). The resulting expression is $\left\langle \delta_{ep} \right\rangle_{\text{sing}} = \int \mathrm{d}\mathbf{r} \delta_{ep}^{\text{sing}}(\mathbf{r})$, where

$$
\begin{aligned}
\delta_{ep}^{\mathrm{sing}}(\mathbf{r}) = & \frac{\left|\psi^{p}(\mathbf{r})\right|^{2}}{\left\langle \Psi^{\text{RXCHF-fe}} \middle| \Psi^{\text{RXCHF-fe}} \right\rangle} \Bigg(\left|\psi_1^{e}(\mathbf{r})\right|^{2} \left[g(\mathbf{r},\mathbf{r}) \right]^{2} \\
& + \sum_{a}^{N_\alpha} \Bigg[\left|\psi_a^{\alpha}(\mathbf{r})\right|^{2} \left\langle \psi_1^{e} \middle| \left[g(\mathbf{r}_1^{e},\mathbf{r}) \right]^{2} \middle| \psi_1^{e} \right\rangle_{\mathbf{r}_1^{e}} - \psi_1^{e*}(\mathbf{r}) \psi_a^{\alpha}(\mathbf{r}) g(\mathbf{r},\mathbf{r}) \left\langle \psi_a^{\alpha} \middle| g(\mathbf{r}_1^{e},\mathbf{r}) \middle| \psi_1^{e} \right\rangle_{\mathbf{r}_1^{e}} \\
& - \psi_a^{\alpha*}(\mathbf{r}) \psi_1^{e}(\mathbf{r}) g(\mathbf{r},\mathbf{r}) \left\langle \psi_1^{e} \middle| g(\mathbf{r}_1^{e},\mathbf{r}) \middle| \psi_a^{\alpha} \right\rangle_{\mathbf{r}_1^{e}} \Bigg] \\
& + \sum_{a}^{N_\alpha} \sum_{b}^{N_\alpha} \Bigg[\psi_b^{\alpha*}(\mathbf{r}) \psi_a^{\alpha}(\mathbf{r}) \left\langle \psi_1^{e} \psi_a^{\alpha} \middle| g(\mathbf{r}_1^{e},\mathbf{r}) g(\mathbf{r}_2^{e},\mathbf{r}) \middle| \psi_b^{\alpha} \psi_1^{e} \right\rangle_{\mathbf{r}_1^{e},\mathbf{r}_2^{e}} \\
& - \left|\psi_b^{\alpha}(\mathbf{r})\right|^{2} \left\langle \psi_1^{e} \psi_a^{\alpha} \middle| g(\mathbf{r}_1^{e},\mathbf{r}) g(\mathbf{r}_2^{e},\mathbf{r}) \middle| \psi_a^{\alpha} \psi_1^{e} \right\rangle_{\mathbf{r}_1^{e},\mathbf{r}_2^{e}} \Bigg] \Bigg).
\end{aligned}
\tag{F.7}
$$

The triplet contribution is determined by evaluating the integral on the right-hand side of Eq. (7.11) using Eq. (F.4). The resulting expression is $\left\langle \delta_{ep} \right\rangle_{\mathrm{trip}} = \int \mathrm{d}\mathbf{r} \delta_{ep}^{\mathrm{trip}}(\mathbf{r})$, where

$$
\begin{aligned}
\left\langle \delta_{ep} \right\rangle_{\mathrm{trip}} = & \frac{\left|\psi^{p}(\mathbf{r})\right|^{2}}{\left\langle \Psi^{\text{RXCHF-fe}} \middle| \Psi^{\text{RXCHF-fe}} \right\rangle} \Bigg(\sum_{a}^{N_\beta} \Bigg[\left|\psi_a^{\beta}(\mathbf{r})\right|^{2} \left\langle \psi_1^{e} \middle| \left[g(\mathbf{r}_1^{e},\mathbf{r}) \right]^{2} \middle| \psi_1^{e} \right\rangle_{\mathbf{r}_1^{e}} \\
& - \sum_{a}^{N_\alpha} \sum_{b}^{N_\beta} \left|\psi_b^{\beta}(\mathbf{r})\right|^{2} \left\langle \psi_1^{e} \psi_a^{\alpha} \middle| g(\mathbf{r}_1^{e},\mathbf{r}) g(\mathbf{r}_2^{e},\mathbf{r}) \middle| \psi_a^{\alpha} \psi_1^{e} \right\rangle_{\mathbf{r}_1^{e},\mathbf{r}_2^{e}} \Bigg).
\end{aligned}
\tag{F.8}
$$

F.3. Positronium Hydride Calculations

The PsH system consists of two electrons, one positron, and a proton. It has previously been studied with the NEO-XCHF,[1,2] as well as the ECG and SVM approaches.[3-6] For a two-electron system with the electrons of opposite spin, the wavefunctions associated with the RXCHF-fe and RXCHF-ne methods are identical, and the RXCHF-ae method is inapplicable. The physical difference between the RXCHF and XCHF wavefunctions is that in the former, only one electronic orbital is explicitly correlated to the positronic orbital, while in the latter, both electronic orbitals are correlated to the positronic orbital. We used the electronic and positronic Gaussian basis sets developed in Ref. 2, where 6 *s* functions were used for both

electronic and positronic basis sets. The XCHF geminal parameters were also obtained from Ref. 2. The RXCHF method with GTG scheme (i) involved optimization of the scaling parameter, which was found to be $\beta = 1.118$ for this system.

The annihilation rate is easily calculated for the RXCHF wavefunction associated with the PsH system. The action of the spin projection operator $\hat{O}^{s}_{e_1 p}$ is to select only combinations where the positronic and first electronic spin coordinates are opposite each other. Assuming an electronically closed-shell system, as the RXCHF methods remove the degeneracy between the two occupied electronic orbitals (i.e., one is geminal-coupled and one is not), we have two equally likely scenarios leading to annihilation. Without loss of generality, we assume that the spin of the positron is α. Either the spin of the geminal-coupled orbital is α and that of the uncoupled orbital is β, or vice versa. The positron can annihilate with the electron occupying the regular (uncoupled) orbital in the first case and with the electron occupying the geminal-coupled orbital in the second case. For RXCHF calculations, we average over the rates corresponding to these two processes.

The energies, average contact densities, and annihilation rates are reported in Table S1. Note that the RXCHF energies are not as low as the XCHF energy, and both are higher than the SVM energy. The objective of the RXCHF method is to describe the local electron-positron interaction accurately, even to the detriment of the overall wavefunction and therefore the total energy. For this system, however, the annihilation rates appear to be slightly less accurate at the RXCHF level than at the XCHF level, although the difference is less than 10% for both RXCHF calculations.

The electron, positron, and electron-positron densities are plotted in Figure S1 for the HF, XCHF, and RXCHF methods with GTG schemes (i) and (ii). The single-particle densities (blue

and red curves) and the contact density (green curves) are more accurate for RXCHF than for XCHF using GTG scheme (i) for both methods. However, the annihilation rate, which is proportional to the integration of the green curves, is less accurate for the RXCHF method than for the XCHF method. The densities in Figure 1b suggest that the highly accurate XCHF annihilation rate arises from a combination of overestimation (underestimation) of the XCHF contact density at distances less than (greater than) $r \approx 3$ a.u.

For the RXCHF method with GTG scheme (ii), the single particle densities agree very well up to $r \approx 2$ a.u., but errors are observed for larger distances, with a shoulder forming in the electronic density and a shifting of the positronic density. These errors cause a shift in the maximum of the contact density to larger distances, but the integration of the contact density is reasonably well-preserved, as evident by the reasonable annihilation rates. GTG scheme (ii), which uses fixed e^-e^+X parameters, is likely more suitable for systems with more positronium character, as these calculations indicate enhanced positronium character by the shift in electronic and positronic densities. Furthermore, comparison of the maxima of the electronic single-particle densities (blue curves) among the different methods illustrates that the shift of density to higher distances in panel (d) actually improves the electronic description of the more tightly bound core electron, which is not explicitly correlated to the positron in the RXCHF calculations.

The comparison of the RXCHF and XCHF methods with GTG scheme (i) shows that RXCHF is able to capture the densities over the entire range, while the XCHF annihilation rates are likely accurate only due to a cancellation of errors. The RXCHF method with the more transferable GTG scheme (ii) does not provide as accurate densities, but the annihilation rate is still within 5% of the accurate SVM value. The PsH system is somewhat unusual in the small degree of positronium character and lack of core electrons, and the RXCHF method is not

expected to be suitable for this special case. The RXCHF method is designed for the larger positron-containing systems studied in Chapter 7.

F.4. Supporting Figures

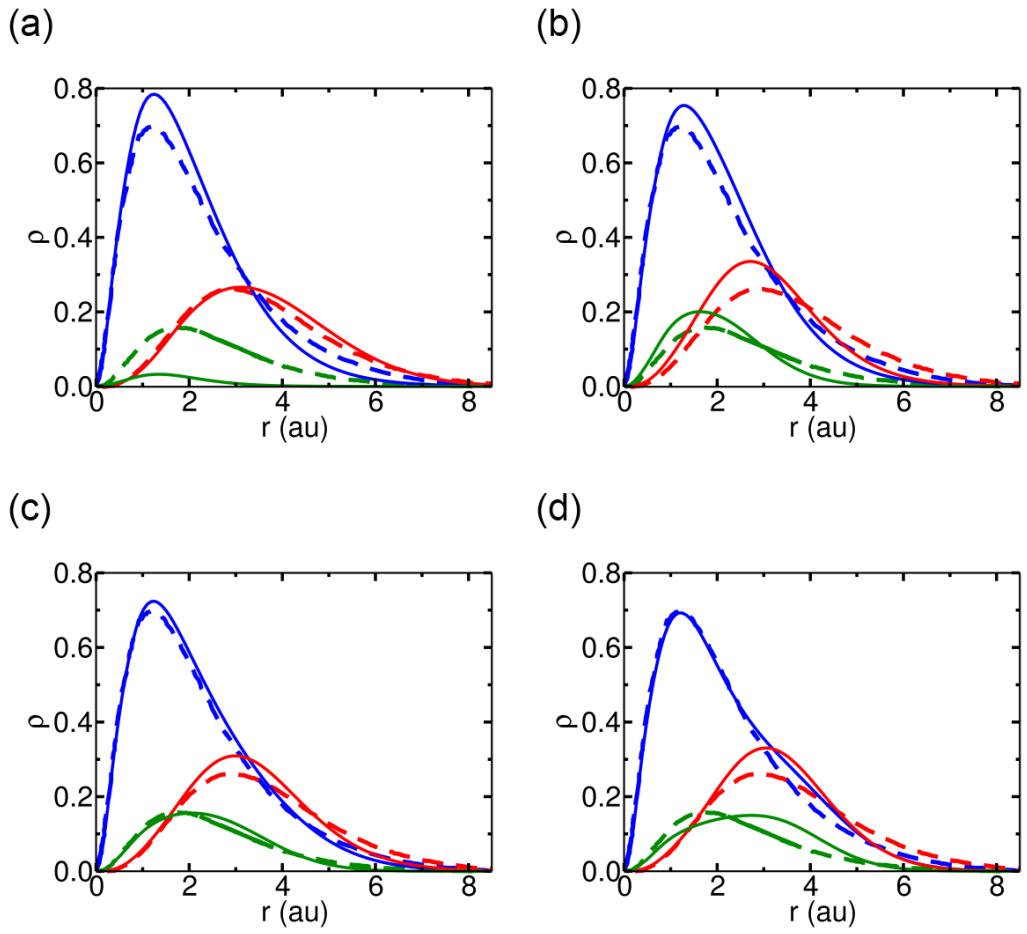

Figure F.1: Radially-averaged electronic (blue) and positronic (red) single-particle densities and contact densities (green) calculated for the PsH system using (a) NEO-HF, (b) XCHF with GTG scheme (i), (c) RXCHF with GTG scheme (i), and (d) RXCHF with GTG scheme (ii) using an even-tempered 6s/6s basis set. The dashed lines represent benchmark ECG data digitized from Ref. 6. The contact densities have been multiplied by 10 for clarity.

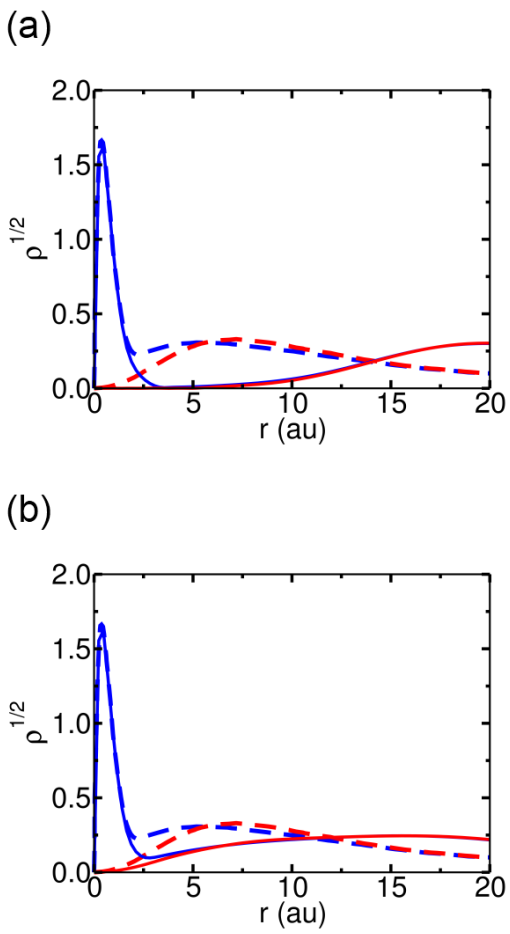

Figure F.2: Radially-averaged electronic (blue) and positronic (red) single-particle densities calculated for the e^{+}Li system using (a) XCHF and (b) RXCHF-fe with GTG schemes (i) using an even-tempered 10s/8s basis set. The dashed lines represent benchmark SVM data digitized from Ref. 5. The analogous densities obtained with GTG scheme (ii) are given in Figure 7.1, and the RXCHF-fe densities with GTG scheme (ii) are in much better agreement with the SVM densities.

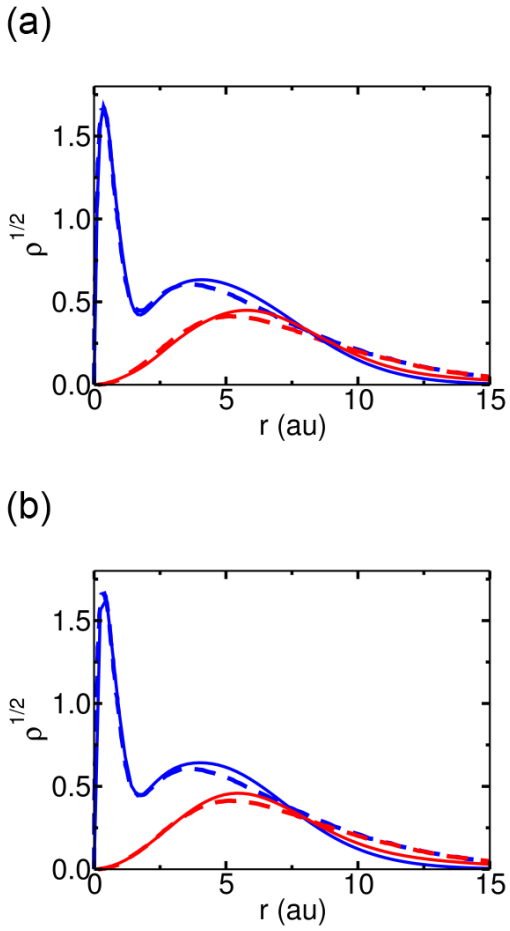

Figure F.3: Radially-averaged electronic (blue) and positronic (red) single-particle densities calculated for the LiPs system using (a) XCHF and (b) RXCHF-ae with GTG scheme (i) using an even-tempered 8s/6s basis set. The dashed lines represent benchmark SVM data digitized from Ref. 5. The analogous densities obtained with GTG scheme (ii) are given in Figure 7.3 and are qualitatively similar.

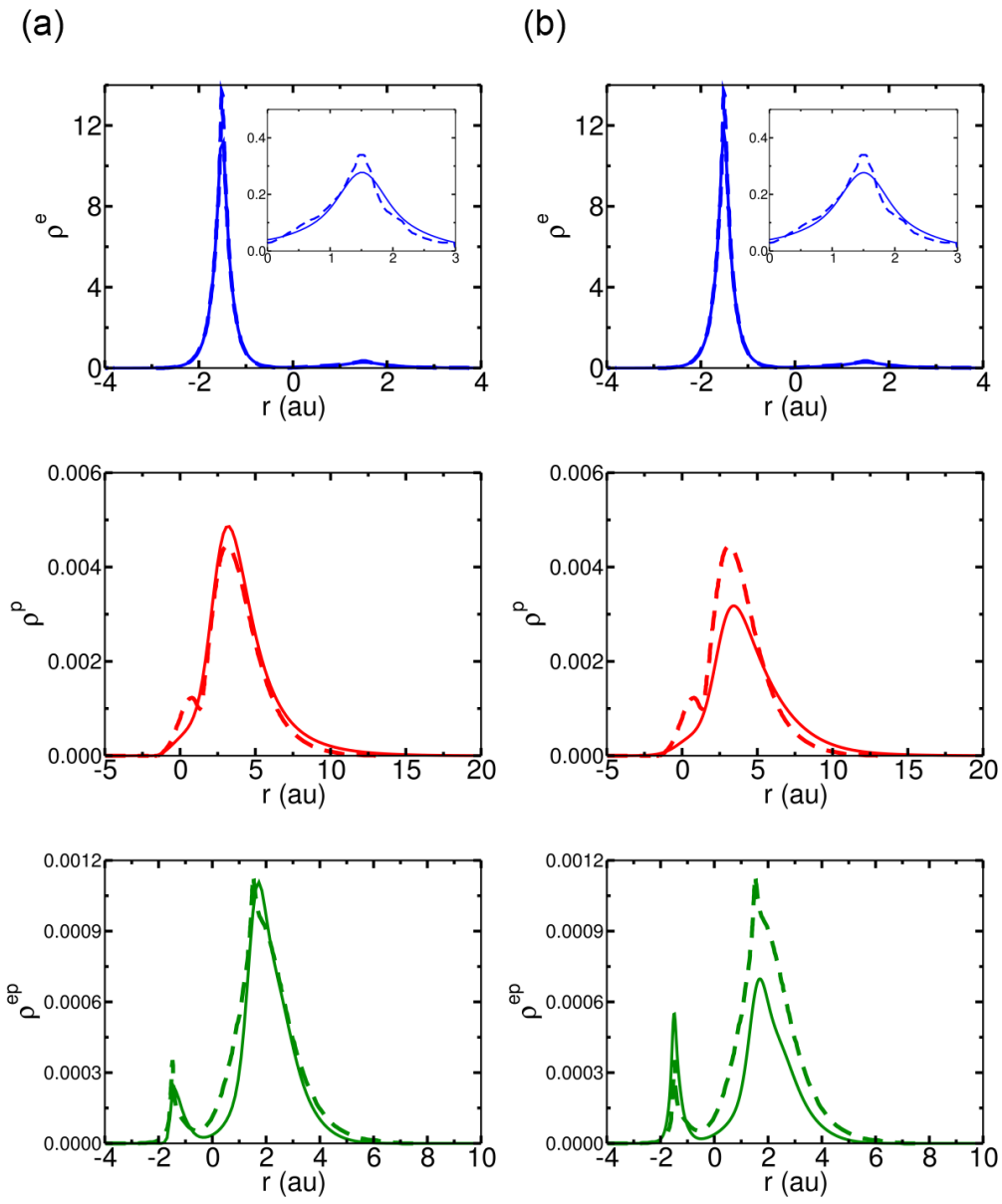

Figure F.4: Electronic (top, blue), positronic (middle, red) and contact (bottom, green) densities calculated for the e^+LiH system using (a) XCHF and (b) RXCHF-ae with GTG scheme (i) using the pc0/6s basis set. The dashed lines represent benchmark ECG data digitized from Ref. 7. The analogous densities obtained with GTG scheme (ii) are given in Figure 7.4, and the RXCHF-ae densities with GTG scheme (ii) are in better agreement with the ECG densities.

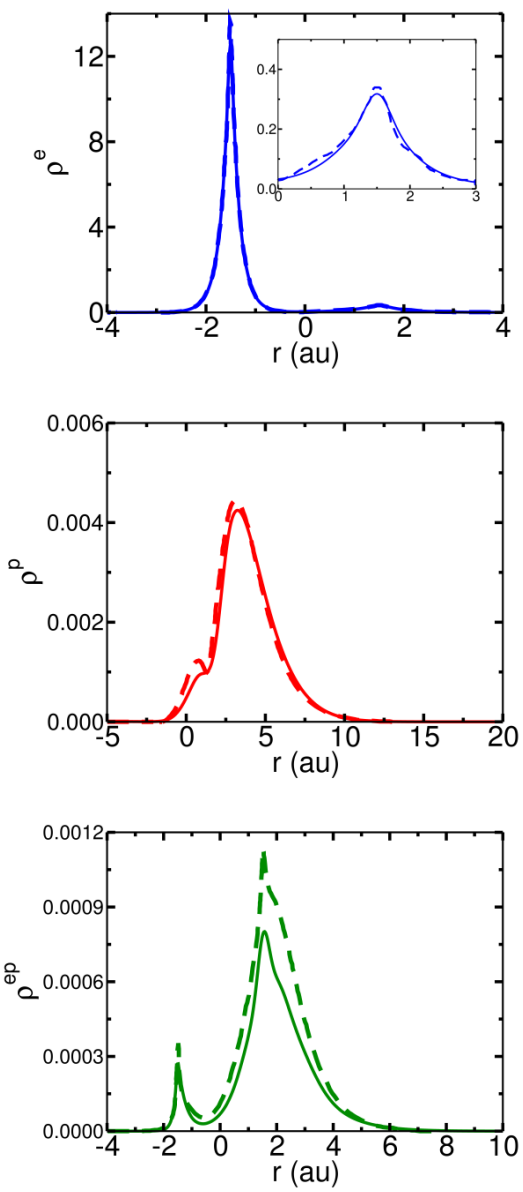

Figure F.5: Electronic (top), positronic (middle) and contact (bottom) densities calculated for the e^+LiH system using RXCHF-ne with GTG scheme (ii) using the pc1/6s-opt basis set. The dashed lines represent benchmark ECG data digitized from Ref. 7. The analogous densities obtained using RXCHF-ae are given in Figure 7.4c and are very similar.

F.5. Supporting Tables

Method (scheme)	Energy	$\langle \delta_{ep} \rangle$	$\Gamma_{2\gamma}$
NEO-HF	-0.66434	0.0063210	0.31902
XCHF (i)	-0.73750	0.049282	2.4872
RXCHF (i)	-0.71532	0.047273	2.3858
RXCHF (ii)	-0.71224	0.051863	2.6175
SVM	-0.78920	0.048975	2.4718

Table F.1: Quantities calculated for the positronium hydride system using an even-tempered 6s/6s basis set with the GTG scheme indicated in the first column. All quantities are given in atomic units except annihilation rates given in ns^{-1}. The SVM data, as well as the prefactor of 100.9394 used to calculate annihilation rates, were obtained from Ref. 4.

Method	Energy	$\langle \delta_{ep} \rangle$	Γ_a	Γ_s	Γ_t
NEO-HF	-7.4260	6.0150×10^{-6}	3.3556×10^{-4}	1.1503×10^{-3}	6.3975×10^{-5}
XCHF	-7.4616	0.081246	4.1006	16.402	2.0612×10^{-5}
RXCHF-ne	-7.4646	0.080284	4.0521	16.207	4.5840×10^{-4}
RXCHF-ae	-7.4648	0.080109	4.0436	16.172	7.3820×10^{-4}
RXCHF-fe	-7.4648	0.080109	4.0434	16.172	7.1514×10^{-4}
SVM	-7.5323	0.034698	1.7512	6.9956	3.0833×10^{-3}

Table F.2: Quantities calculated for the positron-lithium system using an even-tempered 10s/8s basis set with GTG scheme (i) for the XCHF and RXCHF methods. All quantities are given in atomic units except annihilation rates given in ns^{-1}. The SVM data, as well as the prefactor of 201.8788 used to calculate annihilation rates, were obtained from Ref. 8. The analogous data obtained with GTG scheme (ii) are given in Table 7.2 and are in much better agreement with the SVM data.

Method	Energy	$\langle \delta_{ep} \rangle$	Γ_a	Γ_s	Γ_t
RXCHF-ne	-7.4812	0.033092	1.6709	6.6805	1.0283×10^{-3}
RXCHF-ae	-7.4818	0.032946	1.6638	6.6489	2.0980×10^{-3}
RXCHF-fe	-7.4818	0.032946	1.6638	6.6490	2.0631×10^{-3}
SVM	-7.5323	0.034698	1.7512	6.9956	3.0833×10^{-3}

Table F.3: Quantities calculated for the positron-lithium system using an even-tempered 12s/10s basis set with GTG scheme (ii) for the RXCHF methods. All quantities are given in atomic units except annihilation rates given in ns^{-1}. The SVM data, as well as the prefactor to calculate annihilation rates, was taken from Ref. 8 and was 201.8788. The analogous data obtained with the 10s/8s basis set are given in Table 7.2, and the agreement between the two data sets indicates convergence of the basis set.

Method	Energy	$\langle \delta_{ep} \rangle$	$\Gamma_{2\gamma}$
NEO-HF	-7.5260	0.0010167	0.051150
XCHF	-7.6122	0.032096	1.6147
RXCHF-ne	-7.6263	0.071649	3.6046
RXCHF-ae	-7.6286	0.070542	3.5489
SVM	-7.7386	0.04188	2.107

Table F.4: Quantities calculated for the lithium positride system using an even-tempered 8s/6s basis set with GTG scheme (i) for the XCHF and RXCHF methods. All quantities are given in atomic units except annihilation rates given in ns^{-1}. The SVM data, as well as the prefactor of 100.6174809 used to calculate annihilation rates, were obtained from Ref. 5. The analogous data obtained with GTG scheme (ii) are given in Table 7.3 and are in much better agreement with the SVM data.

Method	Energy	$\langle \delta_{ep} \rangle$	$\Gamma_{2\gamma}$
NEO-HF	-7.9674	0.00061346	0.030961
XCHF	-7.9834	0.012685	0.64023
RXCHF-ne	-7.9714	0.0099581	0.50209
RXCHF-ae	-7.9774	0.0093660	0.47224
RXCHF-ne*	-7.9839	0.0077935	0.39295
RXCHF-ae*	-7.9915	0.0074184	0.37404
ECG	-8.1049	0.02499	1.26

Table F.5: Quantities calculated for the positron-lithium hydride system with GTG scheme (i) for the XCHF and RXCHF methods. All quantities are given in atomic units except annihilation rates given in ns^{-1}. The ECG data, as well as the prefactor of 100.8403 used to calculate annihilation rates, were obtained from Ref. 7. All values were calculated with the pc0/6s basis set except those with an asterisk, which were calculated with the pc1/6s-opt basis set. The analogous data obtained with GTG scheme (ii) are given in Table 7.4 and are in much better agreement with the ECG data.

References

(1) M. V. Pak, A. Chakraborty, and S. Hammes-Schiffer, *Journal of Physical Chemistry A* **113**, 4004 (2009).
(2) C. Swalina, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Physics* **136**, 164105 (2012).
(3) S. Bubin and L. Adamowicz, *Physical Review A* **74**, 052502 (2006).
(4) J. Mitroy, *Physical Review A* **73**, 054502 (2006).
(5) G. G. Ryzhikh, J. Mitroy, and K. Varga, *Journal of Physics B-Atomic Molecular and Optical Physics* **31**, 3965 (1998).
(6) K. Strasburger and M. Wolcyrz, *Molecular Physics* **105**, 467 (2007).
(7) K. Strasburger, *Journal of Chemical Physics* **111**, 10555 (1999).
(8) J. Mitroy, *Physical Review A* **70**, 024502 (2004).

Appendix G

Supporting Information for Chapter 8‡

G.1. Summary of Contents

In this Appendix, we include supporting information corresponding to the material presented in Chapter 8. We include the Fock expressions in the spin orbital, spatial orbital, and atomic orbital bases for RXCHF-ae and RXCHF-ne. We also include the symmetric forms of the RXCHF operators. In addition, we include the HF and XCHF contributions to the RXCHF energy, as well as a more compact form for the RXCHF-ne Fock operators. We also include the following figure: electronic molecular orbitals for HCN calculated using NEO-HF and RXCHF-ae.

G.2. Fock operator expressions

In this section, we report expressions for the Fock operators in the spin-orbital, spatial-orbital, and atomic-orbital bases for the RXCHF-ae method, noting that the corresponding equations for the RXCHF-ne method can be generated by setting all "exchange" operators, $\left\{\Omega_2^{(\mathrm{ex})}, \Omega_3^{(\mathrm{ex},1)}, \Omega_3^{(\mathrm{ex},2)}\right\}$, to zero. As discussed in Chapter 8, we report only quantities contributing to the interaction terms between the HF and XCHF subsystems, as the terms for each subsystem

are available from other sources,[1,2] or easily derived from the analogous energy expressions given in Section G.4. The operators defined in this section are all expressed in their symmetrized forms denoted by a tilde. The symmetrization procedure was described in detail in Appendix F, and the symmetric forms of the operators used in this study are given in Section G.2.

1. Spin-orbital basis

The interaction contribution to the regular electronic Fock operator is obtained by varying the energy in Eq. (8.23) with respect to some $\chi_\mu^r,\ \mu \in \{1,\ldots,N_r\}$, leading to

$$\begin{aligned} f^{r,\mathrm{int}}\left(\mathbf{x}_1^r\right) = \frac{1}{S^{\text{RXCHF-ne}}}\Bigg(& \sum_{a=1}^{N_s}\Big[\left\langle \chi^p\chi_a^s\left|\tilde{\Omega}_2\right|\chi^p\chi_a^s\right\rangle - \left\langle \chi^p\chi_a^s\left|\tilde{\Omega}_2^{(\mathrm{ex})}P_2\left(1,2\right)\right|\chi^p\chi_a^s\right\rangle\Big] \\ & + \sum_{a=1}^{N_s}\sum_{b=1}^{N_s}\Big[\left\langle \chi^p\chi_a^s\chi_b^s\left|\tilde{\Omega}_3\right|\chi^p\chi_a^s\chi_b^s\right\rangle - \left\langle \chi^p\chi_a^s\chi_b^s\left|\tilde{\Omega}_3^{(1)}\right|\chi^p\chi_b^s\chi_a^s\right\rangle \\ & - \left\langle \chi^p\chi_a^s\chi_b^s\left|\tilde{\Omega}_3^{(\mathrm{ex},1)}P_2\left(1,2\right)\right|\chi^p\chi_a^s\chi_b^s\right\rangle + \left\langle \chi^p\chi_a^s\chi_b^s\left|\tilde{\Omega}_3^{(\mathrm{ex},1)}P_2\left(1,2\right)\right|\chi^p\chi_b^s\chi_a^s\right\rangle \\ & - \left\langle \chi^p\chi_a^s\chi_b^s\left|\tilde{\Omega}_3^{(\mathrm{ex},2)}P_2\left(1,3\right)\right|\chi^p\chi_a^s\chi_b^s\right\rangle + \left\langle \chi^p\chi_a^s\chi_b^s\left|\tilde{\Omega}_3^{(\mathrm{ex},2)}P_2\left(1,3\right)\right|\chi^p\chi_b^s\chi_a^s\right\rangle\Big] \\ & + \sum_{a=1}^{N_s}\sum_{b=1}^{N_s}\sum_{c=1}^{N_s}\left\langle \chi^p\chi_a^s\chi_b^s\chi_c^s\left|\tilde{\Omega}_4\sum_{m=1}^{3!}(-1)^{p_m}\,P_m^{(3)}\left(2,3,4\right)\right|\chi^p\chi_a^s\chi_b^s\chi_c^s\right\rangle\Bigg). \end{aligned} \tag{G.1}$$

In this expression and those that follow, $P_2\left(i,j\right)$ is the second-order permutation that exchanges electronic indices i and j, and when explicit dependence of the $\tilde{\Omega}_i^{(j)}$ operators is omitted, we assume that the dependent variables are ordered as $p,1,2,\ldots$, e.g., $\tilde{\Omega}_2^{(1)} \equiv \tilde{\Omega}_2^{(1)}\left(p,1,2\right)$.

The interaction contribution to the special electronic Fock operator is obtained by varying the energy in Eq. (8.23) with respect to some $\chi_\mu^s,\ \mu \in \{1,\ldots,N_s\}$. As discussed in Appendix F, compact expressions for $f^{q,\mathrm{int}}$ for RXCHF-ae are not easily obtainable due to the omission of

certain exchange contributions and the ensuing difficulties with operator symmetrizations. Instead, assuming real spin orbitals, we express $f^{s,\mathrm{int}}$ in the spin orbital basis as

$$\begin{aligned}
\left\langle \chi_\mu^s \middle| f^{s,\mathrm{int}}\left(\mathbf{x}_2^s\right) \middle| \chi_\nu^s \right\rangle = \frac{1}{S^{\text{RXCHF-ne}}} \Bigg(& \sum_{a=1}^{N_r} \Big[\left\langle \chi^p \chi_a^r \chi_\mu^s \middle| \tilde{\Omega}_2 \middle| \chi^p \chi_a^r \chi_\nu^s \right\rangle - \left\langle \chi^p \chi_a^r \chi_\mu^s \middle| \tilde{\Omega}_2^{(\mathrm{ex})} \middle| \chi^p \chi_\nu^s \chi_a^r \right\rangle \Big] \\
& + \sum_{a=1}^{N_r} \sum_{b=1}^{N_s} \Big[2 \left\langle \chi^p \chi_a^r \chi_\mu^s \chi_b^s \middle| \tilde{\Omega}_3^{(1)} \middle| \chi^p \chi_a^r \chi_\nu^s \chi_b^s \right\rangle - 2 \left\langle \chi^p \chi_a^r \chi_\mu^s \chi_b^s \middle| \tilde{\Omega}_3^{(1)} \middle| \chi^p \chi_a^r \chi_b^s \chi_\nu^s \right\rangle \\
& - \left\langle \chi^p \chi_a^r \chi_\mu^s \chi_b^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},1)} \middle| \chi^p \chi_\nu^s \chi_a^r \chi_b^s \right\rangle + \left\langle \chi^p \chi_a^r \chi_\mu^s \chi_b^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},1)} \middle| \chi^p \chi_b^s \chi_a^r \chi_\nu^s \right\rangle \\
& - \left\langle \chi^p \chi_a^r \chi_\mu^s \chi_b^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},2)} \middle| \chi^p \chi_b^s \chi_\nu^s \chi_a^r \right\rangle + \left\langle \chi^p \chi_a^r \chi_\mu^s \chi_b^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},2)} \middle| \chi^p \chi_\nu^s \chi_b^s \chi_a^r \right\rangle \\
& - \left\langle \chi^p \chi_a^r \chi_b^s \chi_\mu^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},1)} \middle| \chi^p \chi_b^s \chi_a^r \chi_\nu^s \right\rangle + \left\langle \chi^p \chi_a^r \chi_b^s \chi_\mu^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},1)} \middle| \chi^p \chi_\nu^s \chi_a^r \chi_b^s \right\rangle \\
& - \left\langle \chi^p \chi_a^r \chi_b^s \chi_\mu^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},2)} \middle| \chi^p \chi_\nu^s \chi_b^s \chi_a^r \right\rangle + \left\langle \chi^p \chi_a^r \chi_b^s \chi_\mu^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},2)} \middle| \chi^p \chi_b^s \chi_\nu^s \chi_a^r \right\rangle \Big] \\
& 3 \sum_{a=1}^{N_r} \sum_{b=1}^{N_s} \sum_{c=1}^{N_s} \left\langle \chi^p \chi_a^r \chi_\mu^s \chi_b^s \chi_c^s \middle| \tilde{\Omega}_4 \sum_{m=1}^{3!} (-1)^{p_m} P_m^{(3)}(2,3,4) \middle| \chi^p \chi_a^r \chi_\nu^s \chi_b^s \chi_c^s \right\rangle \Bigg) \\
& - \frac{E^{\mathrm{int-ae}}}{S^{\text{RXCHF-ne}}} \Big(\left\langle \chi^p \chi_\mu^s \middle| g^2 \middle| \chi^p \chi_\nu^s \right\rangle \\
& + 2 \sum_{a=1}^{N_s} \Big[\left\langle \chi^p \chi_a^s \chi_\mu^s \middle| g(1,p)\, g(2,p) \middle| \chi^p \chi_a^s \chi_\nu^s \right\rangle \\
& - \left\langle \chi^p \chi_a^s \chi_\mu^s \middle| g(1,p)\, g(2,p) \middle| \chi^p \chi_\nu^s \chi_a^s \right\rangle \Big] \Big).
\end{aligned} \tag{G.2}$$

for any spin orbitals χ_μ^s, χ_ν^s with $\mu,\nu \in \{1,\ldots N_s\}$. We note that the compact form of $f^{s,\mathrm{int}}$ can be readily obtained for the RXCHF-ne level by removal of all terms in Eq. (G.2) involving exchange operators and the introduction of P_2 operators. This form is included in Section G.5.

The interaction contribution to the quantum nuclear Fock operator is obtained by varying the energy in Eq. (8.23) with respect to χ^p, leading to

$$
\begin{aligned}
f^{\,p,\mathrm{int}}\left(\mathbf{x}^{p}\right) = \frac{1}{S^{\mathrm{RXCHF\text{-}ne}}}\Bigg(& \sum_{a=1}^{N_r}\sum_{b=1}^{N_s}\Big[\left\langle \chi_a^r \chi_b^s \left| \tilde{\Omega}_2 \right| \chi_a^r \chi_b^s \right\rangle - \left\langle \chi_a^r \chi_b^s \left| \tilde{\Omega}_2^{(\mathrm{ex})} \right| \chi_b^s \chi_a^r \right\rangle\Big] \\
& + \sum_{a=1}^{N_r}\sum_{b=1}^{N_s}\sum_{c=1}^{N_s}\Big[\left\langle \chi_a^r \chi_b^s \chi_c^s \left| \tilde{\Omega}_3 \right| \chi_a^r \chi_b^s \chi_c^s \right\rangle - \left\langle \chi_a^r \chi_b^s \chi_c^s \left| \tilde{\Omega}_3 \right| \chi_a^r \chi_c^s \chi_b^s \right\rangle \\
& - \left\langle \chi_a^r \chi_b^s \chi_c^s \left| \tilde{\Omega}_3^{(\mathrm{ex},1)} \right| \chi_b^s \chi_a^r \chi_c^s \right\rangle + \left\langle \chi_a^r \chi_b^s \chi_c^s \left| \tilde{\Omega}_3^{(\mathrm{ex},1)} \right| \chi_c^s \chi_a^r \chi_b^s \right\rangle \\
& - \left\langle \chi_a^r \chi_b^s \chi_c^s \left| \tilde{\Omega}_3^{(\mathrm{ex},2)} \right| \chi_c^s \chi_b^s \chi_a^r \right\rangle + \left\langle \chi_a^r \chi_b^s \chi_c^s \left| \tilde{\Omega}_3^{(\mathrm{ex},2)} \right| \chi_b^s \chi_c^s \chi_a^a \right\rangle\Big] \\
& + \sum_{a=1}^{N_r}\sum_{b=1}^{N_s}\sum_{c=1}^{N_s}\sum_{d=1}^{N_s}\left\langle \chi_a^r \chi_b^s \chi_c^s \chi_d^s \left| \tilde{\Omega}_4 \sum_{m=1}^{3!}(-1)^{p_m} P_m^{(3)}(2,3,4) \right| \chi_a^r \chi_b^s \chi_c^s \chi_d^s \right\rangle\Bigg) \\
& - \frac{E^{\mathrm{int-ae}}}{S^{\mathrm{RXCHF\text{-}ne}}}\Bigg(\sum_{a=1}^{N_s}\left\langle \chi_a^s \left| g^2 \right| \chi_a^s \right\rangle \\
& + \sum_{a=1}^{N_s}\sum_{b=1}^{N_s}\Big[\left\langle \chi_a^s \chi_b^s \left| g(1,p)\, g(2,p) \right| \chi_a^s \chi_b^s \right\rangle - \left\langle \chi_a^s \chi_b^s \left| g(1,p)\, g(2,p) \right| \chi_b^s \chi_a^s \right\rangle\Big]\Bigg).
\end{aligned}
\tag{G.3}
$$

2. Spatial-orbital basis

Next we integrate the RXCHF-ae Fock operators in Eqs. (G.1) to (G.3) over spin to obtain the spatial Fock operators assuming a closed-shell subsystem for both regular and special electrons (i.e., using Eq. (8.27) and its analogue for the special electron orbitals). The interaction contribution to the regular electronic Fock operator is given by

$$
\begin{aligned}
f^{r,\mathrm{int}}\left(\mathbf{r}_1^r\right) = \frac{1}{S^{\mathrm{RXCHF\text{-}ne}}}\Bigg(& \sum_{a=1}^{N_s/2}\Big[2\left\langle \psi^p\psi_a^s \middle| \tilde{\Omega}_2 \middle| \psi^p\psi_a^s \right\rangle - \left\langle \psi^p\psi_a^s \middle| \tilde{\Omega}_2^{(\mathrm{ex})} P_2(1,2) \middle| \psi^p\psi_a^s \right\rangle \Big] \\
& + \sum_{a=1}^{N_s/2}\sum_{b=1}^{N_s/2}\Big[4\left\langle \psi^p\psi_a^s\psi_b^s \middle| \tilde{\Omega}_3 \middle| \psi^p\psi_a^s\psi_b^s \right\rangle - 2\left\langle \psi^p\psi_a^s\psi_b^s \middle| \tilde{\Omega}_3 \middle| \psi^p\psi_b^s\psi_a^s \right\rangle \\
& - 2\left\langle \psi^p\psi_a^s\psi_b^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},1)} P_2(1,2) \middle| \psi^p\psi_a^s\psi_b^s \right\rangle + \left\langle \psi^p\psi_a^s\psi_b^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},1)} P_2(1,2) \middle| \psi^p\psi_b^s\psi_a^s \right\rangle \\
& - 2\left\langle \psi^p\psi_a^s\psi_b^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},2)} P_2(1,3) \middle| \psi^p\psi_a^s\psi_b^s \right\rangle + \left\langle \psi^p\psi_a^s\psi_b^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},2)} P_2(1,3) \middle| \psi^p\psi_b^s\psi_a^s \right\rangle \Big] \\
& + 2\sum_{a=1}^{N_s/2}\sum_{b=1}^{N_s/2}\sum_{c=1}^{N_s/2} \\
& \Big[4\left\langle \psi^p\psi_a^s\psi_b^s\psi_c^s \middle| \tilde{\Omega}_4 \middle| \psi^p\psi_a^s\psi_b^s\psi_c^s \right\rangle - 2\left\langle \psi^p\psi_a^s\psi_b^s\psi_c^s \middle| \tilde{\Omega}_4 \middle| \psi^p\psi_a^s\psi_c^s\psi_b^s \right\rangle \\
& - 2\left\langle \psi^p\psi_a^s\psi_b^s\psi_c^s \middle| \tilde{\Omega}_4 \middle| \psi^p\psi_b^s\psi_a^s\psi_c^s \right\rangle + \left\langle \psi^p\psi_a^s\psi_b^s\psi_c^s \middle| \tilde{\Omega}_4 \middle| \psi^p\psi_b^s\psi_c^s\psi_a^s \right\rangle \\
& + \left\langle \psi^p\psi_a^s\psi_b^s\psi_c^s \middle| \tilde{\Omega}_4 \middle| \psi^p\psi_c^s\psi_a^s\psi_b^s \right\rangle - 2\left\langle \psi^p\psi_a^s\psi_b^s\psi_c^s \middle| \tilde{\Omega}_4 \middle| \psi^p\psi_c^s\psi_b^s\psi_a^s \right\rangle \Big] \Bigg).
\end{aligned}
\tag{G.4}
$$

The interaction contribution to the special electronic Fock operator is given by

$$\left\langle \psi_\mu^s \middle| f^{s,\mathrm{int}}\left(\mathbf{r}_2^s\right) \middle| \psi_\nu^s \right\rangle = \frac{1}{S^{\mathrm{RXCHF\text{-}ne}}} \left(\sum_{a=1}^{N_r/2} \left[2\left\langle \psi^p \psi_a^r \psi_\mu^s \middle| \tilde{\Omega}_2 \middle| \psi^p \psi_a^r \psi_\nu^s \right\rangle - \left\langle \psi^p \psi_a^r \psi_\mu^s \middle| \tilde{\Omega}_2^{(\mathrm{ex})} \middle| \psi^p \psi_\nu^s \psi_a^r \right\rangle \right] \right.$$

$$+ \sum_{a=1}^{N_r/2} \sum_{b=1}^{N_s/2} [$$

$$8\left\langle \psi^p \psi_a^r \psi_\mu^s \psi_b^s \middle| \tilde{\Omega}_3^{(1)} \middle| \psi^p \psi_a^r \psi_\nu^s \psi_b^s \right\rangle - 4\left\langle \psi^p \psi_a^r \psi_\mu^s \psi_b^s \middle| \tilde{\Omega}_3^{(1)} \middle| \psi^p \psi_a^r \psi_b^s \psi_\nu^s \right\rangle$$

$$- \left\langle \psi^p \psi_a^r \psi_\mu^s \psi_b^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},1)} \middle| \psi^p \psi_\nu^s \psi_a^r \psi_b^s \right\rangle + \left\langle \psi^p \psi_a^r \psi_\mu^s \psi_b^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},1)} \middle| \psi^p \psi_b^s \psi_a^r \psi_\nu^s \right\rangle$$

$$- \left\langle \psi^p \psi_a^r \psi_\mu^s \psi_b^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},2)} \middle| \psi^p \psi_b^s \psi_\nu^s \psi_a^r \right\rangle + \left\langle \psi^p \psi_a^r \psi_\mu^s \psi_b^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},2)} \middle| \psi^p \chi_\nu^s \psi_b^s \psi_a^r \right\rangle$$

$$- \left\langle \psi^p \psi_a^r \psi_b^s \psi_\mu^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},1)} \middle| \psi^p \psi_b^s \psi_a^r \psi_\nu^s \right\rangle + \left\langle \psi^p \psi_a^r \psi_b^s \psi_\mu^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},1)} \middle| \psi^p \psi_\nu^s \psi_a^r \psi_b^s \right\rangle$$

$$- \left\langle \psi^p \psi_a^r \psi_b^s \psi_\mu^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},2)} \middle| \psi^p \chi_\nu^s \psi_b^s \psi_a^r \right\rangle + \left\langle \psi^p \psi_a^r \psi_b^s \psi_\mu^s \middle| \tilde{\Omega}_3^{(\mathrm{ex},2)} \middle| \psi^p \psi_b^s \psi_\nu^s \psi_a^r \right\rangle \Big]$$

$$+ 6 \sum_{a=1}^{N_r/2} \sum_{b=1}^{N_s/2} \sum_{c=1}^{N_s/2}$$

$$\Big[4\left\langle \psi^p \psi_a^r \psi_\mu^s \psi_b^s \psi_c^s \middle| \tilde{\Omega}_4 \middle| \psi^p \psi_a^r \psi_\nu^s \psi_b^s \psi_c^s \right\rangle - 2\left\langle \psi^p \psi_a^r \psi_\mu^s \psi_b^s \psi_c^s \middle| \tilde{\Omega}_4 \middle| \psi^p \psi_a^r \psi_\nu^s \psi_c^s \psi_b^s \right\rangle$$

$$- 2\left\langle \psi^p \psi_a^r \psi_\mu^s \psi_b^s \psi_c^s \middle| \tilde{\Omega}_4 \middle| \psi^p \psi_a^r \psi_b^s \psi_\nu^s \psi_c^s \right\rangle + \left\langle \psi^p \psi_a^r \psi_\mu^s \psi_b^s \psi_c^s \middle| \tilde{\Omega}_4 \middle| \psi^p \psi_a^r \psi_b^s \psi_c^s \psi_\nu^s \right\rangle$$

$$+ \left\langle \psi^p \psi_a^r \psi_\mu^s \psi_b^s \psi_c^s \middle| \tilde{\Omega}_4 \middle| \psi^p \psi_a^r \psi_c^s \psi_\nu^s \psi_b^s \right\rangle - 2\left\langle \psi^p \psi_a^r \psi_\mu^s \psi_b^s \psi_c^s \middle| \tilde{\Omega}_4 \middle| \psi^p \psi_a^r \psi_c^s \psi_b^s \psi_\nu^s \right\rangle \Big] \Big)$$

$$- \frac{E^{\mathrm{int-ae}}}{S^{\mathrm{RXCHF\text{-}ne}}} \Big(\left\langle \psi^p \psi_\mu^s \middle| g^2 \middle| \psi^p \psi_\nu^s \right\rangle \qquad \text{(G.5)}$$

$$+ 2 \sum_{a=1}^{N_s/2} \Big[2\left\langle \psi^p \psi_a^s \psi_\mu^s \middle| g(1,p)\, g(2,p) \middle| \psi^p \psi_a^s \psi_\nu^s \right\rangle$$

$$- \left\langle \psi^p \psi_a^s \psi_\mu^s \middle| g(1,p)\, g(2,p) \middle| \psi^p \psi_\nu^s \psi_a^s \right\rangle \Big] \Big).$$

The interaction contribution to the quantum nuclear Fock operator is given by

$$
\begin{aligned}
f^{p,\mathrm{int}}\left(\mathbf{r}^{p}\right) = \frac{1}{S^{\text{RXCHF-ne}}}\Bigg(& 2\sum_{a=1}^{N_r/2}\sum_{b=1}^{N_s/2}\Big[2\left\langle\psi_a^r\psi_b^s\left|\tilde{\Omega}_2\right|\psi_a^r\psi_b^s\right\rangle - \left\langle\psi_a^r\psi_b^s\left|\tilde{\Omega}_2^{(\mathrm{ex})}\right|\psi_b^s\psi_a^r\right\rangle\Big] \\
& +2\sum_{a=1}^{N_r/2}\sum_{b=1}^{N_s/2}\sum_{c=1}^{N_s/2}\Big[4\left\langle\psi_a^r\psi_b^s\psi_c^s\left|\tilde{\Omega}_3\right|\psi_a^r\psi_b^s\psi_c^s\right\rangle - 2\left\langle\psi_a^r\psi_b^s\psi_c^s\left|\tilde{\Omega}_3\right|\psi_a^r\psi_c^s\psi_b^s\right\rangle \\
& -2\left\langle\psi_a^r\psi_b^s\psi_c^s\left|\tilde{\Omega}_3^{(\mathrm{ex},1)}\right|\psi_b^s\psi_a^r\psi_c^s\right\rangle + \left\langle\psi_a^r\psi_b^s\psi_c^s\left|\tilde{\Omega}_3^{(\mathrm{ex},1)}\right|\psi_c^s\psi_a^r\psi_b^s\right\rangle \\
& -2\left\langle\psi_a^r\psi_b^s\psi_c^s\left|\tilde{\Omega}_3^{(\mathrm{ex},2)}\right|\psi_c^s\psi_b^s\psi_a^r\right\rangle + \left\langle\psi_a^r\psi_b^s\psi_c^s\left|\tilde{\Omega}_3^{(\mathrm{ex},2)}\right|\psi_b^s\psi_c^s\psi_a^a\right\rangle\Big] \\
& +4\sum_{a=1}^{N_r/2}\sum_{b=1}^{N_s/2}\sum_{c=1}^{N_s/2}\sum_{d=1}^{N_s/2} \\
& \Big[4\left\langle\psi_a^r\psi_b^s\psi_c^s\psi_d^s\left|\tilde{\Omega}_4\right|\psi_a^r\psi_b^s\psi_c^s\psi_d^s\right\rangle - 2\left\langle\psi_a^r\psi_b^s\psi_c^s\psi_d^s\left|\tilde{\Omega}_4\right|\psi_a^r\psi_b^s\psi_d^s\psi_c^s\right\rangle \\
& -2\left\langle\psi_a^r\psi_b^s\psi_c^s\psi_d^s\left|\tilde{\Omega}_4\right|\psi_a^r\psi_c^s\psi_b^s\psi_d^s\right\rangle + \left\langle\psi_a^r\psi_b^s\psi_c^s\psi_d^s\left|\tilde{\Omega}_4\right|\psi_a^r\psi_c^s\psi_d^s\psi_b^s\right\rangle \\
& +\left\langle\psi_a^r\psi_b^s\psi_c^s\psi_d^s\left|\tilde{\Omega}_4\right|\psi_a^r\psi_d^s\psi_b^s\psi_c^s\right\rangle - 2\left\langle\psi_a^r\psi_b^s\psi_c^s\psi_d^s\left|\tilde{\Omega}_4\right|\psi_a^r\psi_d^s\psi_c^s\psi_b^s\right\rangle\Big]\Bigg) \\
& -\frac{E^{\mathrm{int-ae}}}{S^{\text{RXCHF-ne}}}\Bigg(2\sum_{a=1}^{N_s/2}\left\langle\psi_a^s\left|g^2\right|\psi_a^s\right\rangle \\
& +2\sum_{a=1}^{N_s/2}\sum_{b=1}^{N_s/2}\Big[2\left\langle\psi_a^s\psi_b^s\left|g(1,p)\,g(2,p)\right|\psi_a^s\psi_b^s\right\rangle - \left\langle\psi_a^s\psi_b^s\left|g(1,p)\,g(2,p)\right|\psi_b^s\psi_a^s\right\rangle\Big]\Bigg).
\end{aligned}
\tag{G.6}
$$

3. Atomic-orbital basis

We now expand the spatial orbitals in the AO bases, as in Eqs. (8.28) to (8.30). We first define density matrices as

$$
P_{\mu\nu}^{r} = \sum_{i=1}^{N_r/2} C_{\mu,\mathrm{i}}^{\mathrm{r}^*} C_{\nu,\mathrm{i}}^{r}, \tag{G.7}
$$

$$
P_{\mu"\nu"}^{s} = \sum_{i=1}^{N_s/2} C_{\mu",\mathrm{i}}^{\mathrm{s}^*} C_{\nu",\mathrm{i}}^{s}, \tag{G.8}
$$

$$
P_{\mu'\nu'}^{p} = C_{\mu'}^{p^*} C_{\nu'}^{p}. \tag{G.9}
$$

The interaction contribution to the regular electronic Fock operator in the AO basis is given by

$$
\begin{aligned}
F^{r,\mathrm{int}}_{\mu\nu} = & \frac{1}{S^{\text{RXCHF-ne}}} \\
& \Bigg(\sum_{\mu'\nu'} \sum_{\mu"\nu"} P^{p}_{\mu'\nu'} P^{s}_{\mu"\nu"} \Big[\Gamma_2\left(\mu',\nu';\mu,\nu;\mu",\nu"\right) - \frac{1}{2}\Gamma^{(\mathrm{ex})}_2\left(\mu',\nu';\mu,\nu;\mu",\nu"\right) \Big] \\
& + \sum_{\mu'\nu'} \sum_{\mu_1"\nu_1"} \sum_{\mu_2"\nu_2"} P^{p}_{\mu'\nu'} P^{s}_{\mu_1"\nu_1"} P^{s}_{\mu_2"\nu_2"} \Big[\Gamma_3\left(\mu',\nu';\mu,\nu;\mu_1",\nu_1";\mu_2",\nu_2"\right) \\
& - \frac{1}{2}\Gamma^{(\mathrm{ex},1)}_3\left(\mu',\nu';\mu,\nu;\mu_1",\nu_1";\mu_2",\nu_2"\right) - \frac{1}{2}\Gamma^{(\mathrm{ex},2)}_3\left(\mu',\nu';\mu,\nu;\mu_1",\nu_1";\mu_2",\nu_2"\right) \Big] \\
& + \sum_{\mu'\nu'} \sum_{\mu_1"\nu_1"} \sum_{\mu_2"\nu_2"} \sum_{\mu_3",\nu_3"} P^{p}_{\mu'\nu'} P^{s}_{\mu_1"\nu_1"} P^{s}_{\mu_2"\nu_2"} P^{s}_{\mu_3",\nu_3"} \Gamma_4\left(\mu',\nu';\mu,\nu;\mu_1",\nu_1";\mu_2",\nu_2";\mu_3",\nu_3"\right) \Bigg).
\end{aligned}
\tag{G.10}
$$

The interaction contribution to the special electronic Fock operator in the AO basis is given by

$$
\begin{aligned}
F^{s}_{\mu_1"\nu_1"} = & \frac{1}{S^{\text{RXCHF-ne}}} \\
& \Bigg(\sum_{\mu\nu} \sum_{\mu'\nu'} P^{r}_{\mu\nu} P^{p}_{\mu'\nu'} \Big[\Gamma_2\left(\mu',\nu';\mu,\nu;\mu_1",\nu_1"\right) - \frac{1}{2}\Gamma^{(\mathrm{ex})}_2\left(\mu',\nu';\mu,\nu;\mu_1",\nu_1"\right) \Big] \\
& + \sum_{\mu\nu} \sum_{\mu'\nu'} \sum_{\mu_2"\nu_2"} P^{r}_{\mu\nu} P^{p}_{\mu'\nu'} P^{s}_{\mu_2"\nu_2"} \Big[2\Gamma_3\left(\mu',\nu';\mu,\nu;\mu_1",\nu_1";\mu_2",\nu_2"\right) \\
& - \frac{1}{2}\Gamma^{(\mathrm{ex},1)}_3\left(\mu',\nu';\mu,\nu;\mu_1",\nu_1";\mu_2",\nu_2"\right) - \frac{1}{2}\Gamma^{(\mathrm{ex},2)}_3\left(\mu',\nu';\mu,\nu;\mu_1",\nu_1";\mu_2",\nu_2"\right) \\
& - \frac{1}{2}\Gamma^{(\mathrm{ex},1)}_3\left(\mu',\nu';\mu,\nu;\mu_2",\nu_2";\mu_1",\nu_1"\right) - \frac{1}{2}\Gamma^{(\mathrm{ex},2)}_3\left(\mu',\nu';\mu,\nu;\mu_2",\nu_2";\mu_1",\nu_1"\right) \Big] \\
& + 3\sum_{\mu\nu} \sum_{\mu'\nu'} \sum_{\mu_2"\nu_2"} \sum_{\mu_3",\nu_3"} P^{r}_{\mu\nu} P^{p}_{\mu'\nu'} P^{s}_{\mu_2"\nu_2"} P^{s}_{\mu_3",\nu_3"} \Gamma_4\left(\mu',\nu';\mu,\nu;\mu_1",\nu_1";\mu_2",\nu_2";\mu_3",\nu_3"\right) \Bigg) \\
& - \frac{E^{\mathrm{int-ae}}}{S^{\text{RXCHF-ne}}} \Bigg(\sum_{\mu'\nu'} P^{p}_{\mu'\nu'} \Gamma_{1s}\left(\mu',\nu';\mu_1",\nu_1"\right) \\
& + 2\sum_{\mu'\nu'} \sum_{\mu_2"\nu_2"} P^{p}_{\mu'\nu'} P^{s}_{\mu_2"\nu_2"} \Gamma_{2s}\left(\mu',\nu';\mu_1",\nu_1";\mu_2",\nu_2"\right) \Bigg).
\end{aligned}
\tag{G.11}
$$

The interaction contribution to the nuclear Fock operator in the AO basis is given by

$$
\begin{aligned}
F^{p,\mathrm{int}}_{\mu'\nu'} = & \frac{1}{S^{\text{RXCHF-ne}}} \\
& \Bigg(\sum_{\mu\nu} \sum_{\mu"\nu"} P^{r}_{\mu\nu} P^{s}_{\mu"\nu"} \left[\Gamma_2\left(\mu',\nu';\mu,\nu;\mu",\nu"\right) - \frac{1}{2}\Gamma^{(\mathrm{ex})}_2\left(\mu',\nu';\mu,\nu;\mu",\nu"\right) \right] \\
& + \sum_{\mu\nu} \sum_{\mu_1"\nu_1"} \sum_{\mu_2"\nu_2"} P^{r}_{\mu\nu} P^{s}_{\mu_1"\nu_1"} P^{s}_{\mu_2"\nu_2"} \Big[\Gamma_3\left(\mu',\nu';\mu,\nu;\mu_1",\nu_1";\mu_2",\nu_2"\right) \\
& - \frac{1}{2}\Gamma^{(\mathrm{ex},1)}_3\left(\mu',\nu';\mu,\nu;\mu_1",\nu_1";\mu_2",\nu_2"\right) - \frac{1}{2}\Gamma^{(\mathrm{ex},2)}_3\left(\mu',\nu';\mu,\nu;\mu_1",\nu_1";\mu_2",\nu_2"\right) \Big] \\
& + \sum_{\mu\nu} \sum_{\mu_1"\nu_1"} \sum_{\mu_2"\nu_2"} \sum_{\mu_3",\nu_3"} P^{r}_{\mu\nu} P^{s}_{\mu_1"\nu_1"} P^{s}_{\mu_2"\nu_2"} P^{s}_{\mu_3",\nu_3"} \Gamma_4\left(\mu',\nu';\mu,\nu;\mu_1",\nu_1";\mu_2",\nu_2";\mu_3",\nu_3"\right) \Bigg) \\
& - \frac{E^{\mathrm{int-ae}}}{S^{\text{RXCHF-ne}}} \Bigg(\sum_{\mu"\nu"} P^{s}_{\mu"\nu"} \Gamma_{1s}\left(\mu',\nu';\mu",\nu"\right) \\
& + \sum_{\mu_1"\nu_1"} \sum_{\mu_2"\nu_2"} P^{s}_{\mu_1"\nu_1"} P^{s}_{\mu_2"\nu_2"} \Gamma_{2s}\left(\mu',\nu';\mu_1",\nu_1";\mu_2",\nu_2"\right) \Bigg).
\end{aligned}
\tag{G.12}
$$

The integrals required for the evaluation of these quantities are defined as

$$
\Gamma_{1s}\left(\mu',\nu';\mu",\nu"\right) = \left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu"} \middle| g^2 \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu"} \right\rangle,
\tag{G.13}
$$

$$
\Gamma_2\left(\mu',\nu';\mu,\nu;\mu",\nu"\right) = \left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu}\phi^{e}_{\mu"} \middle| \tilde{\Omega}_2 \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu}\phi^{e}_{\nu"} \right\rangle,
\tag{G.14}
$$

$$
\Gamma^{(\mathrm{ex})}_2\left(\mu',\nu';\mu",\nu";\mu,\nu\right) = \left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu}\phi^{e}_{\mu"} \middle| \tilde{\Omega}^{(\mathrm{ex})}_2 \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu"}\phi^{e}_{\nu} \right\rangle,
\tag{G.15}
$$

$$
\begin{aligned}
\Gamma_{2s}\left(\mu',\nu';\mu_1",\nu_1";\mu_2",\nu_2"\right) = & \left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu_1"}\phi^{e}_{\mu_2"} \middle| g(1)\,g(2) \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu_1"}\phi^{e}_{\nu_2"} \right\rangle \\
& - \frac{1}{2}\left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu_1"}\phi^{e}_{\mu_2"} \middle| g(1)\,g(2) \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu_2"}\phi^{e}_{\nu_1"} \right\rangle,
\end{aligned}
\tag{G.16}
$$

$$
\begin{aligned}
\Gamma_3\left(\mu',\nu';\mu,\nu;\mu_1",\nu_1";\mu_2",\nu_2"\right) = & \left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu}\phi^{e}_{\mu_1"}\phi^{e}_{\mu_2"} \middle| \tilde{\Omega}_3 \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu}\phi^{e}_{\nu_1"}\phi^{e}_{\nu_2"} \right\rangle \\
& - \frac{1}{2}\left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu}\phi^{e}_{\mu_1"}\phi^{e}_{\mu_2"} \middle| \tilde{\Omega}_3 \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu}\phi^{e}_{\nu_2"}\phi^{e}_{\nu_1"} \right\rangle,
\end{aligned}
\tag{G.17}
$$

$$
\begin{aligned}
\Gamma^{(\mathrm{ex},1)}_3\left(\mu',\nu';\mu,\nu;\mu_1",\nu_1";\mu_2",\nu_2"\right) = & \left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu}\phi^{e}_{\mu_1"}\phi^{e}_{\mu_2"} \middle| \tilde{\Omega}^{(\mathrm{ex},1)}_3 \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu_1"}\phi^{e}_{\nu}\phi^{e}_{\nu_2"} \right\rangle \\
& - \frac{1}{2}\left\langle \phi^{p}_{\mu'}\phi^{e}_{\mu}\phi^{e}_{\mu_1"}\phi^{e}_{\mu_2"} \middle| \tilde{\Omega}^{(\mathrm{ex},1)}_3 \middle| \phi^{p}_{\nu'}\phi^{e}_{\nu_2"}\phi^{e}_{\nu}\phi^{e}_{\nu_1"} \right\rangle,
\end{aligned}
\tag{G.18}
$$

$$\Gamma_3^{(\mathrm{ex},2)}\left(\mu',\nu';\mu,\nu;\mu_1'',\nu_1'';\mu_2'',\nu_2''\right)=\left\langle\phi_\mu^p\phi_\mu^e\phi_{\mu_1''}^e\phi_{\mu_2''}^e\left|\tilde{\Omega}_3^{(\mathrm{ex},2)}\right|\phi_\nu^p\phi_{\nu_2''}^e\phi_{\nu_1''}^e\phi_\nu^e\right\rangle -\frac{1}{2}\left\langle\phi_\mu^p\phi_\mu^e\phi_{\mu_1''}^e\phi_{\mu_2''}^e\left|\tilde{\Omega}_3^{(\mathrm{ex},2)}\right|\phi_\nu^p\phi_{\nu_1''}^e\phi_{\nu_2''}^e\phi_\nu^e\right\rangle, \tag{G.19}$$

$$\begin{aligned}\Gamma_4\left(\mu',\nu';\mu,\nu;\mu_1'',\nu_1'';\mu_2'',\nu_2'';\mu_3'',\nu_3''\right)=&\left\langle\phi_\mu^p\phi_\mu^e\phi_{\mu_1''}^e\phi_{\mu_2''}^e\phi_{\mu_3''}^e\left|\tilde{\Omega}_4\right|\phi_\nu^p\phi_\nu^e\phi_{\nu_1''}^e\phi_{\nu_2''}^e\phi_{\nu_3''}^e\right\rangle\\ &-\frac{1}{2}\left\langle\phi_\mu^p\phi_\mu^e\phi_{\mu_1''}^e\phi_{\mu_2''}^e\phi_{\mu_3''}^e\left|\tilde{\Omega}_4\right|\phi_\nu^p\phi_\nu^e\phi_{\nu_1''}^e\phi_{\nu_3''}^e\phi_{\nu_2''}^e\right\rangle\\ &-\frac{1}{2}\left\langle\phi_\mu^p\phi_\mu^e\phi_{\mu_1''}^e\phi_{\mu_2''}^e\phi_{\mu_3''}^e\left|\tilde{\Omega}_4\right|\phi_\nu^p\phi_\nu^e\phi_{\nu_2''}^e\phi_{\nu_1''}^e\phi_{\nu_3''}^e\right\rangle\\ &+\frac{1}{4}\left\langle\phi_\mu^p\phi_\mu^e\phi_{\mu_1''}^e\phi_{\mu_2''}^e\phi_{\mu_3''}^e\left|\tilde{\Omega}_4\right|\phi_\nu^p\phi_\nu^e\phi_{\nu_2''}^e\phi_{\nu_3''}^e\phi_{\nu_1''}^e\right\rangle\\ &+\frac{1}{4}\left\langle\phi_\mu^p\phi_\mu^e\phi_{\mu_1''}^e\phi_{\mu_2''}^e\phi_{\mu_3''}^e\left|\tilde{\Omega}_4\right|\phi_\nu^p\phi_\nu^e\phi_{\nu_3''}^e\phi_{\nu_1''}^e\phi_{\nu_2''}^e\right\rangle\\ &-\frac{1}{2}\left\langle\phi_\mu^p\phi_\mu^e\phi_{\mu_1''}^e\phi_{\mu_2''}^e\phi_{\mu_3''}^e\left|\tilde{\Omega}_4\right|\phi_\nu^p\phi_\nu^e\phi_{\nu_3''}^e\phi_{\nu_2''}^e\phi_{\nu_1''}^e\right\rangle.\end{aligned} \tag{G.20}$$

G.3. Symmetric Operator Forms

In this section, we report the symmetric forms of the operators used in the working expressions developed in Section G.2. Usage of the symmetric forms facilitates the straightforward Fock procedures described in Chapter 8. More detail about the symmetrization procedure can be found in Appendix F. Herein, we simply report the appropriate symmetric forms of the RXCHF operators used in this study:

$$\tilde{\Omega}_2\left(p,1,2\right)=\Omega_2\left(p,1,2\right), \tag{G.21}$$

$$\tilde{\Omega}_2^{(\mathrm{ex})}\left(p,1,2\right)=\frac{1}{2}\left(\Omega_2^{(\mathrm{ex})}\left(p,1,2\right)+\Omega_2^{(\mathrm{ex})}\left(p,2,1\right)\right), \tag{G.22}$$

$$\tilde{\Omega}_3\left(p,1,2,3\right)=\frac{1}{2}\left(\Omega_3\left(p,1,2,3\right)+\Omega_3\left(p,1,3,2\right)\right), \tag{G.23}$$

$$\tilde{\Omega}_3^{(\mathrm{ex},1)}\left(p,1,2,3\right)=\frac{1}{2}\left(\Omega_3^{(\mathrm{ex},1)}\left(p,1,2,3\right)+\Omega_3^{(\mathrm{ex},1)}\left(p,2,1,3\right)\right), \tag{G.24}$$

$$\tilde{\Omega}_3^{(\mathrm{ex},2)}\left(p,1,2,3\right)=\frac{1}{2}\left(\Omega_3^{(\mathrm{ex},2)}\left(p,1,2,3\right)+\Omega_3^{(\mathrm{ex},2)}\left(p,3,2,1\right)\right), \tag{G.25}$$

$$\tilde{\Omega}_4\left(p,1,2,3,4\right)=\frac{1}{3!}\sum_{m=1}^{3!}P_m^{(3)}\left(2,3,4\right)\Omega_4\left(p,1,2,3,4\right). \tag{G.26}$$

G.4. HF and XCHF Energy Contributions

In Chapter 8, the RXCHF-ne and RXCHF-ae energy expressions were decomposed as $E=E^{\mathrm{HF}}+E^{\mathrm{XCHF}}+E^{\mathrm{int}}$, where detailed expressions for the interaction energy, E^{int}, were derived in detail. In what follows, we report analogous expressions for E^{HF} and E^{XCHF}. These expressions can be used to derive analogs of the Fock expressions for the interaction energy that were given in Section G.2.

The Hartree-Fock energy, E^{HF}, corresponds to the energy of an N_r-electron system and can be found in standard quantum chemistry textbooks.[2] In terms of the regular electron spin orbitals, $\left\{\chi_a^r\right\}$, defined in the main text, E^{HF} is given by

$$\mathrm{E}^{\mathrm{HF}}=\sum_{a=1}^{N_r}\left\langle \chi_a^r\left|h^e\right|\chi_a^r\right\rangle+\sum_{a=1}^{N_r}\sum_{b=1}^{N_r}\left(\left\langle \chi_a^r\chi_b^r\left|\frac{1}{r_{12}}\right|\chi_a^r\chi_b^r\right\rangle-\left\langle \chi_a^r\chi_b^r\left|\frac{1}{r_{12}}\right|\chi_b^r\chi_a^r\right\rangle\right), \tag{G.27}$$

where h^e is the standard one-electron operator and the angular brackets and ordering of spin orbitals (with respect to coordinate dependences) follow the convention outlined in Chapter 8.

The NEO-XCHF energy, E^{XCHF}, corresponds to the explicitly-correlated Hartree-Fock energy of system of N_s electrons and 1 quantum proton. As described in Chapter 8, this can be obtained from Ref. 1 by modifying the operators therein through removal of all terms with zero

or one geminal factor, g. In terms of the special electron spin orbitals, $\{\chi_a^s\}$, and the quantum proton spin orbital, χ^p, defined in the main text, the resulting energy expression is given by

$$\mathrm{E}^{\mathrm{XCHF}} = \frac{E_1 + E_2 + E_3 + E_4}{S^{\mathrm{XCHF}}}. \tag{G.28}$$

In this expression, $S^{\mathrm{XCHF}} = S^{\mathrm{RXCHF\text{-}ne}}$ which is given in Eq. (8.11), and

$$\begin{aligned}
E_1 &= \sum_{a=1}^{N_s} \left\langle \chi^p \chi_a^s \left| \Omega_1^{\mathrm{XCHF}} \right| \chi^p \chi_a^s \right\rangle, \\
E_2 &= \sum_{a=1}^{N_s} \sum_{b=1}^{N_s} \left\langle \chi^p \chi_a^s \chi_b^s \left| \Omega_2^{\mathrm{XCHF}} \sum_{m=1}^{2!} (-1)^{p_m} \, P_m^{(2)} \right| \chi^p \chi_a^s \chi_b^s \right\rangle, \\
E_3 &= \sum_{a=1}^{N_s} \sum_{b=1}^{N_s} \sum_{c=1}^{N_s} \left\langle \chi^p \chi_a^s \chi_b^s \chi_c^s \left| \Omega_3^{\mathrm{XCHF}} \sum_{m=1}^{3!} (-1)^{p_m} \, P_m^{(3)} \right| \chi^p \chi_a^s \chi_b^s \chi_c^s \right\rangle, \\
E_4 &= \sum_{a=1}^{N_s} \sum_{b=1}^{N_s} \sum_{c=1}^{N_s} \sum_{d=1}^{N_s} \left\langle \chi^p \chi_a^s \chi_b^s \chi_c^s \chi_d^s \left| \Omega_4^{\mathrm{XCHF}} \sum_{m=1}^{4!} (-1)^{p_m} \, P_m^{(4)} \right| \chi^p \chi_a^s \chi_b^s \chi_c^s \chi_d^s \right\rangle,
\end{aligned} \tag{G.29}$$

with the XCHF operators given by

$$\begin{aligned}
\Omega_1^{\mathrm{XCHF}}(p,1) &= g(p,1)\left[h^p(p) + h^e(1) + V_{ep}(p,1) \right] g(p,1), \\
\Omega_2^{\mathrm{XCHF}}(p,1,2) &= g(p,2)\left[h^e(1) + V_{ep}(p,1) + V_{ee}(1,2) \right] g(p,2) \\
&\quad + g(p,1)\left[h^p(p) + h^e(1) + h^e(2) + 2V_{ep}(p,1) + V_{ee}(1,2) \right] g(p,2), \\
\Omega_3^{\mathrm{XCHF}}(p,1,2,3) &= g(p,3)\left[V_{ee}(1,2) \right] g(p,3) \\
&\quad + g(p,2)\left[h^e(1) + V_{ep}(p,1) + 2V_{ee}(1,2) \right] g(p,3), \\
\Omega_4^{\mathrm{XCHF}}(p,1,2,3,4) &= g(p,3)\left[V_{ee}(1,2) \right] g(p,4),
\end{aligned} \tag{G.30}$$

where h^p is the standard one-proton operator and the other quantities are defined in Chapter 8.

G.5. RXCHF-ne Special Electronic Fock Operators

In Section G.2, the interaction contribution to the Fock operator for the special electron orbitals was derived for RXCHF-ae and was reported in terms of matrix elements over either

spin or spatial orbitals. As noted, a compact form for this operator can be written for RXCHF-ne as there are no complications arising from symmetrizing the extra exchange terms appearing for RXCHF-ae.

The interaction contribution to the special electronic Fock operator in the spin orbital basis is given by

$$
\begin{aligned}
f^{\mathrm{s,int}}\left(\mathbf{x}_2^s\right) = \frac{1}{S^{\mathrm{RXCHF\text{-}ne}}}\Bigg(& \sum_{a=1}^{N_r}\left\langle \chi^p \chi_a^r \left| \tilde{\Omega}_2 \right| \chi^p \chi_a^r \right\rangle \\
& + 2\sum_{a=1}^{N_r}\sum_{b=1}^{N_s}\left[\left\langle \chi^p \chi_a^r \chi_b^s \left| \tilde{\Omega}_3 \right| \chi^p \chi_a^r \chi_b^s \right\rangle - \left\langle \chi^p \chi_a^r \chi_b^s \left| \tilde{\Omega}_3 P_2\left(2,3\right) \right| \chi^p \chi_a^r \chi_b^s \right\rangle \right] \\
& + 3\sum_{a=1}^{N_r}\sum_{b=1}^{N_s}\sum_{c=1}^{N_s}\left\langle \chi^p \chi_a^r \chi_b^s \chi_c^s \left| \tilde{\Omega}_4 \sum_{m=1}^{3!}\left(-1\right)^{p_m} P_m^{(3)}\left(2,3,4\right) \right| \chi^p \chi_a^r \chi_b^s \chi_c^s \right\rangle \Bigg) \\
& - \frac{E^{\mathrm{int-ne}}}{S^{\mathrm{RXCHF\text{-}ne}}}\Bigg(\left\langle \chi^p \left| g^2 \right| \chi^p \right\rangle \\
& + 2\sum_{a=1}^{N_s}\left[\left\langle \chi^p \chi_a^s \left| g\left(1,p\right) g\left(2,p\right) \right| \chi^p \chi_a^s \right\rangle - \left\langle \chi^p \chi_a^s \left| g\left(1,p\right) g\left(2,p\right) P_2\left(1,2\right) \right| \chi^p \chi_a^s \right\rangle \right] \Bigg),
\end{aligned}
\tag{G.31}
$$

while in the spatial orbital basis, this Fock operator is given by

$$
\begin{aligned}
f^{\mathrm{s,int}}\left(\mathbf{r}_2^s\right) = \frac{1}{S^{\mathrm{RXCHF\text{-}ne}}}\Bigg(& 2\sum_{a=1}^{N_r/2}\left\langle \psi^p \psi_a^r \left| \tilde{\Omega}_2 \right| \psi^p \psi_a^r \right\rangle \\
& + 4\sum_{a=1}^{N_r/2}\sum_{b=1}^{N_s/2}\left[2\left\langle \psi^p \psi_a^r \psi_b^s \left| \tilde{\Omega}_3 \right| \psi^p \psi_a^r \psi_b^s \right\rangle - \left\langle \psi^p \psi_a^r \psi_b^s \left| \tilde{\Omega}_3 P_2\left(2,3\right) \right| \psi^p \psi_a^r \psi_b^s \right\rangle \right] \\
& + 6\sum_{a=1}^{N_r/2}\sum_{b=1}^{N_s/2}\sum_{c=1}^{N_s/2} \\
& \Big[4\left\langle \psi^p \psi_a^r \psi_b^s \psi_c^s \left| \tilde{\Omega}_4 \right| \psi^p \psi_a^r \psi_b^s \psi_c^s \right\rangle - 2\left\langle \psi^p \psi_a^r \psi_b^s \psi_c^s \left| \tilde{\Omega}_4 \right| \psi^p \psi_a^r \psi_c^s \psi_b^s \right\rangle \\
& - 2\left\langle \psi^p \psi_a^r \psi_b^s \psi_c^s \left| \tilde{\Omega}_4 P_2\left(2,3\right) \right| \psi^p \psi_a^r \psi_b^s \psi_c^s \right\rangle + \left\langle \psi^p \psi_a^r \psi_b^s \psi_c^s \left| \tilde{\Omega}_4 P_2\left(2,4\right) \right| \psi^p \psi_a^r \psi_c^s \psi_b^s \right\rangle \\
& + \left\langle \psi^p \psi_a^r \psi_b^s \psi_c^s \left| \tilde{\Omega}_4 P_2\left(2,3\right) \right| \psi^p \psi_a^r \psi_c^s \psi_b^s \right\rangle - 2\left\langle \psi^p \psi_a^r \psi_b^s \psi_c^s \left| \tilde{\Omega}_4 P_2\left(2,4\right) \right| \psi^p \psi_a^r \psi_b^s \psi_c^s \right\rangle \Big] \Bigg) \\
& - \frac{E^{\mathrm{int-ne}}}{S^{\mathrm{RXCHF\text{-}ne}}}\Bigg(\left\langle \psi^p \left| g^2 \right| \chi^p \right\rangle \\
& + 2\sum_{a=1}^{N_s/2}\left[2\left\langle \psi^p \psi_a^s \left| g\left(1,p\right) g\left(2,p\right) \right| \psi^p \psi_a^s \right\rangle - \left\langle \psi^p \psi_a^s \left| g\left(1,p\right) g\left(2,p\right) P_2\left(1,2\right) \right| \psi^p \psi_a^s \right\rangle \right] \Bigg).
\end{aligned}
\tag{G.32}
$$

G.6. Supporting Figures

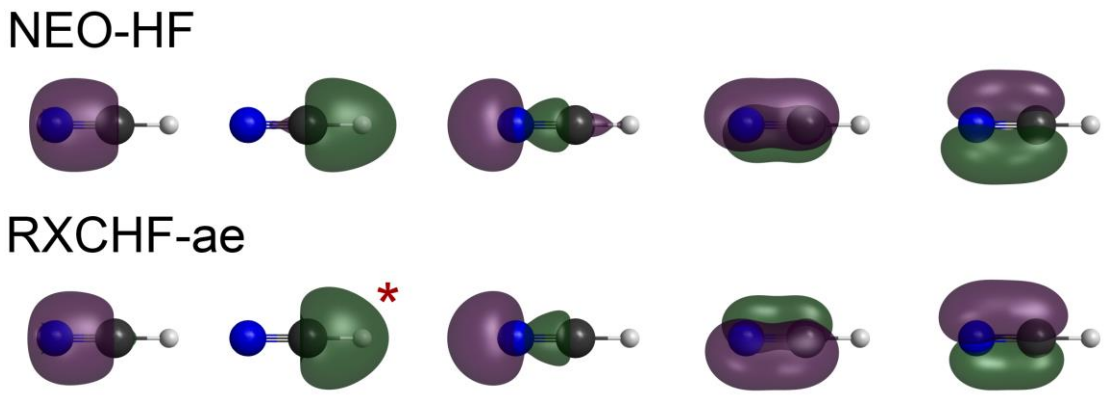

Figure G.1: Electronic molecular orbitals for HCN obtained using different methods, where green and purple denote the different phases of the orbitals. The top row consists of five of the occupied MOs obtained from a converged NEO-HF calculation ordered from left to right according to their respective eigenvalues from lowest to highest. The bottom row consists of four of the occupied regular MOs and the special MO obtained from the converged RXCHF-ae calculation. The regular MOs are ordered with respect to their respective eigenvalues from lowest to highest, but the special MO, denoted with an asterisk, is placed in between them to highlight its similarity to the MOs in the previous row. In both rows, the two occupied MOs that were excluded from this image correspond to the core orbitals on the N and C centers (i.e., the two lowest-eigenvalue orbitals for the NEO-HF and RXCHF-ne calculations).

References

(1) A. Chakraborty, M. V. Pak, and S. Hammes-Schiffer, *Journal of Chemical Physics* **129**, 014101 (2008).

(2) A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. (Dover, New York, 1996).